

10/068426

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: October 14, 2003, 06:16:09 ; Search time 25.3023 Seconds  
(without alignments)  
2067.627 Million cell updates/sec

Title: US-10-068-426-1  
Perfect score: 2906  
Sequence: 1 MPLLALLLPSPHPPIC.....MHEALHHYTKSLSPGK 544  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1693	58.3	626	1 NBHUIA	platelet glycoprotein
2	1205	41.5	374	2 S69339	Ig heavy chain V r
3	1197.5	41.2	330	1 GHU	Ig gamma-1 chain C
4	1197	41.2	255	4 S1856	Ig gamma-1 chain C
5	1150	39.6	234	2 P0207	Ig gamma chain C r
6	1149.5	39.6	326	1 G2HU	Ig gamma-2 chain C
7	1141.5	39.3	377	2 A23511	Ig gamma-3 chain C
8	1139.5	39.2	377	2 A60764	Ig gamma-4 chain C
9	1137	39.1	327	1 G4HU	Ig gamma-4 heavy c
10	1115	38.4	289	1 G3HUI	Ig gamma 2b chain
11	918	31.6	328	2 I47160	Ig gamma 2a chain
12	918	31.6	328	2 I47159	Ig gamma chain C r
13	913	31.4	323	1 GHRB	Ig gamma 4 chain c
14	911	31.3	277	2 I47162	Ig gamma 4 chain c
15	888.5	30.6	328	2 I47158	Ig gamma 1 chain c
16	887	30.5	328	2 I47161	Ig gamma 3 chain c
17	886	30.5	328	2 I47161	Ig gamma 3 chain c
18	851	29.3	470	2 S22080	Ig heavy chain pre
19	843	29.0	329	1 G3MSC	Ig gamma-3 chain C
20	835	28.7	308	2 C30554	Ig heavy chain C r
21	835	28.7	472	2 S31459	Ig gamma-1 chain -
22	834	28.7	444	2 PC4336	monoclonal antibod
23	832	28.6	398	1 G3MSM	Ig gamma-3 chain C
24	831	28.6	326	2 PS6017	Ig gamma-1 chain C
25	828	28.5	333	2 PS0018	Ig gamma-2b chain
26	823	28.3	329	2 S00847	Ig gamma-2c chain
27	820.5	28.2	324	1 G1MS	Ig gamma-1 chain C
28	815.5	28.1	393	1 G1MSM	Ig gamma-1 chain C
29	812	27.9	322	2 PS0019	Ig gamma-2a chain

30	805	27.7	330	1 G2MSA	Ig gamma-2a chain
31	805	27.7	469	2 S37483	Ig gamma-2a chain
32	802	27.6	335	1 G2MSAB	Ig gamma-2a chain
33	800	27.5	399	1 G2MSAM	Ig gamma-2a chain
34	790	27.2	446	2 S40295	Ig gamma-2a chain
35	769.5	26.5	474	1 G2MS11	Ig gamma-2b chain
36	762.5	26.2	327	2 S06611	Ig gamma-2 chain C
37	762.5	26.2	327	2 S06611	Ig gamma-2b chain
38	746	25.7	475	2 S01321	Ig gamma-2b chain
39	704	24.2	180	2 I46732	Ig gamma heavy cha
40	580.5	20.0	249	2 S69340	Ig heavy chain VHI
41	573	19.7	152	2 S14236	Ig gamma-1 chain C
42	568.5	19.6	218	2 A36040	Ig heavy chain V-I
43	409	14.1	572	2 B46529	Ig y heavy chain (
44	359.5	12.4	388	1 EHMS	Ig epsilon chain C
45	358	12.3	453	2 S37768	Ig mu chain C regi

ALIGNMENTS

RESULT 1  
NBHUIA  
platelet glycoprotein Ib alpha chain precursor - human  
N:Alternate names: membrane glycoprotein Ib alpha chain  
N:Contains: glycoocalicin  
C:Species: Homo sapiens (man)  
C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 22-Jun-1999  
C:Accession: A94174; A60435; A94173; S16945; I55355; A27075; A27102  
R:Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hagen, F.S.; Papayannopoulou, T.; Roth, G.J  
Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987  
A>Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembran  
A:Reference number: A94174; MUID:87289655; PMID:3303030  
A:Accession: A94174  
A:Molecule type: mRNA  
A:Residues: 1-826 <LOP>  
A:Cross-references: GB:J02940; NID:gl83499; PIDN:AA5595.1; PID:9306793  
R:Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemenston,  
Thromb. Haemost. 61, 448-453, 1989  
A>Title: Isolation and characterization of human blood platelet mRNA and construction  
d cloning of a GPIb coding cDNA insert  
A:Reference number: A60435; MUID:90020160; PMID:2799758  
A:Accession: A60435  
A:Molecule type: mRNA  
A:Residues: 207-467 <WIC>  
R:Titani, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987  
A>Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet  
A:Reference number: A94173; MUID:87289654; PMID:3497398  
A:Accession: A94173  
A:Molecule type: protein  
A:Residues: 17-315 <TIT>  
R:Hess, D.; Schaller, J.; Rickli, E.E.; Clemenston, K.J.  
Eur. J. Biochem. 199, 389-393, 1991  
A>Title: Identification of the disulphide bonds in human platelet glycoocalicin.  
A:Reference number: S16945; MUID:91301149; PMID:2070794  
A:Accession: S16945  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 224-227;262-270;277-282 <HES>  
R:Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.  
J. Biol. Chem. 267, 10055-10061, 1992  
A>Title: Polymorphism of human glycoprotein Ib alpha results from a variable number o  
ations.  
A:Reference number: I55355; MUID:92250564; PMID:1577776  
A:Accession: I55355  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 412-427 <RES>  
A:Cross-references: GB:S34436; NID:g249176; PIDN:AA22152.1; PID:g249177  
A>Note: variant D  
C:Comment: Glycoprotein Ib (GPIb), a surface membrane protein of platelets, participa  
C:Comment: Platelet activation apparently involves disruption of the macromolecular c

C;Comment: Binding sites for von Willebrand factor and thrombin (the latter site with ur  
C;Comment: Glycocalicin, which is approximately coextensive with the extracellular part  
C;Genetics:  
A;Gene: GDB:GPIBA; GPIB  
A;Cross-references: GDB:118806; OMIM:231200  
A;Map position: 17pter-17p12  
C;Complex: heterodimer with platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem repeat  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MP>  
F:48-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:379-430/Region: proline/threonine-rich 9-residue repeats  
F:502-540/Domain: transmembrane #status predicted <TRM>  
F:541-626/Domain: intracellular #status predicted <INT>  
F:37,175/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:308/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match 58.3%; Score 1693; DB 1; Length 626;  
Best Local Similarity 87.4%; Pred. No. 2.9e-108;  
Matches 333; Conservative 4; Mismatches 18; Indels 26; Gaps 5;

Qy 1 MLLLLLLLLLPSLPHPICEVSVASHLEVNCDKRLTALPPDLKDTILHLSNLLY 60  
Db 1 MLLLLLLLLLPSLPHPICEVSVASHLEVNCDKRLTALPPDLKDTILHLSNLLY 60

Qy 61 TFSLATLMPYTRLTQMLDRCLETKLQVDGTLPLVGLTDLSDHQLSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQMLDRCLETKLQVDGTLPLVGLTDLSDHQLSLPLGQTLPALTV 120

Qy 121 LDVSFNRLTSLPLGALRGLEQLYKGNELKTLPLGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNRLTSLPLGALRGLEQLYKGNELKTLPLGLLTPTPKLEKSLANNLTLP 180

Qy 181 AGLLNGLENLDTLLQNSLYTIIPKFGFSGHLLPFAFLHGNPMLNCEIYFRRWLQDNA 240  
Db 181 AGLLNGLENLDTLLQNSLYTIIPKFGFSGHLLPFAFLHGNPMLNCEIYFRRWLQDNA 240

Qy 241 ENVVYVWQGVVDVKAWSNVASVQCDNSDKPVPYKPGKPTLQDGDGDTLYDYYPEEDT 300  
Db 241 ENVVYVWQGVVDVKAWSNVASVQCDNSDKPVPYKPGKPTLQDGDGDTLYDYYPEEDT 300

Qy 301 EGDKVRATRTVVKPPTKARPHTCP-----PCPA---PEALGAPSVFLFPPK 343  
Db 301 EGDKVRATRTVVKPPTKARPHTCP-----PCPA---PEALGAPSVFLFPPK 343

Qy 344 --PKDTI-----MISRTPEVT 357  
Db 359 WTPNFTLHMSITSPKPKST 379

RESULT 2  
S69339  
Ig heavy chain V region precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
C;Accession: S69339; S72664  
R;Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A;Reference number: S69339; MUID:95262687; PMID:7744049  
A;Accession: S69339  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-374 <KHA>  
A;Cross-references: EMBL:X81695  
R;Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994  
A;Reference number: S72664  
A;Accession: S72664  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-140,'C',142-374 <KH2>  
A;Cross-references: EMBL:X81695  
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 41.5%; Score 1205; DB 2; Length 374;  
Best Local Similarity 86.2%; Pred. No. 4e-75;  
Matches 231; Conservative 5; Mismatches 24; Indels 8; Gaps 3;

Qy 295 DEGDLDLY--DYPEEDTEGKV---RATRTVVKPPTKA--RHCTCPCPAPEALGAPS 336  
Db 107 DPADTATTCYGYGEGYGGYRFRHSWGGTTLTVTSSEPKSCDKTHICPPCPAPEALLGAPS 166

Qy 337 VLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNST 396  
Db 167 VLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNST 226

Qy 397 YRVSVLTVLHODMLNGEYKCKVSNKALPVPDKITISKAKGQPREPOVYITLPPSRBEMT 456  
Db 227 YRVSVLTVLHODMLNGEYKCKVSNKALPVPDKITISKAKGQPREPOVYITLPPSRBEMT 286

Qy 457 KNQVSLTCLVKGFPSPDIATVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQ 516  
Db 287 KNQVSLTCLVKGFPSPDIATVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQ 346

Qy 517 GNVFSCSVMEALHNHYTKQSLSPGK 544  
Db 347 GNVFSCSVMEALHNHYTKQSLSPGK 374

RESULT 3  
GHU  
Ig gamma-1 chain C region - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999  
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146  
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A;Reference number: A93433; MUID:82274238; PMID:6287432  
A;Accession: A93433  
A;Molecule type: DNA  
A;Residues: 1-330 <ELL>  
A;Cross-references: EMBL:Z17370  
A;Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) marker  
A;Note: Lys-330 is removed after translation  
R;Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A;Reference number: S33904  
A;Accession: S36861  
A;Molecule type: DNA  
A;Residues: 2-330 <HAR>  
A;Cross-references: EMBL:Z17370  
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of  
A;Reference number: S33887; MUID:83001943; PMID:6811139  
A;Accession: S33887  
A;Molecule type: DNA  
A;Residues: 88-113;235-330 <TAK>  
A;Cross-references: EMBL:Z17370  
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman  
Biochemistry 9, 3161-3170, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq  
A;Reference number: A90563; MUID:71064024; PMID:5489771  
A;Contents: myeloma protein Bu  
A;Accession: B90563  
A;Molecule type: protein  
A;Residues: 1-96,'R',98-135 <CUN>

A:Note: this sequence has the Gln(3) marker, 97-Arg  
 R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
 Biochemistry 9, 3171-3181, 1970  
 A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequence  
 A:Reference number: A90564; MUID:71064025; PMID:5530842  
 A:Contents: Eu  
 A:Accession: A90564  
 A:Molecule type: protein  
 A:Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,  
 A:Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met  
 R:Ponstingl, H.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
 A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),  
 igen Primärstruktur.  
 A:Reference number: A91668; MUID:77070269; PMID:826475  
 A:Contents: myeloma protein Nie  
 A:Accession: B91668  
 A:Molecule type: protein  
 A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27  
 A:Note: this sequence has the Gln(17) and Gln(1) markers  
 R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
 A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL  
 A:Reference number: A91723; MUID:83289331; PMID:6884994  
 A:Contents: myeloma protein KOL; disulfide bonds  
 A:Accession: A91723  
 A:Molecule type: protein  
 A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH  
 A:Note: this sequence has the Gln(3) and Gln(non-1) markers  
 R:Gall, W.E.; Edelman, G.M., 1970  
 Biochemistry 9, 3188-3196, 1970  
 A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
 A:Reference number: A90565; MUID:71064027; PMID:4923144  
 A:Contents: annotation; disulfide bonds  
 R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
 A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
 enamide cleavage products, and the disulfide bridges.  
 A:Reference number: A91667; MUID:77070267; PMID:1002129  
 A:Contents: annotation; disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG1  
 A:Cross-references: GDB:120085; OMIM:147100  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 99/1; 114/1; 224/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:120-85/Domain: immunoglobulin homology <IM1>  
 F:137-206/Domain: immunoglobulin homology <IM2>  
 F:243-310/Domain: immunoglobulin homology <IM3>  
 F:27-83, 144-204, 250-308/Disulfide bonds: #status experimental  
 F:103/Disulfide bonds: interchain (to light chain) #status experimental  
 F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:180/Binding site: carbohydrate (Aan) (covalent) #status experimental

Query Match 41.2%; Score 1197.5; DB 1; Length 330;  
 Best Local Similarity 73.1%; Pred. No. 1.1e-74;  
 Matches 236; Conservative 11; Mismatches 37; Indels 39; Gaps 6;  
 QY 253 KAMTSNVAQCDNSDKFP-----VYKPGKCPGLGDEGDTLDYDYPE 297  
 DB 16 KSTSGTAALGCLVKDYPEPVTVSWNSGALTSGVHTFPA----VLQSSG---LYSLSSV 68  
 QY 298 ETEGDKVATRTVV-----KFTKARP-----HTCPPCPAEPALGAPSVFLFP 341  
 DB 69 VTPVSSSL-GTQYICNVNHNKPSNTKVKKVPKSCDKTHTCPPCPAPELLGGPSVFLFP 127  
 QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREEQYNSTRYVS 401  
 DB 128 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREEQYNSTRYVS 187

QY 402 VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNQVS 461  
 DB 188 VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVS 247  
 QY 462 LTCLVKGFPYSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 521  
 DB 248 LTCLVKGFPYSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 307  
 QY 522 CSVMHRAHNNHYTKQSLSPGK 544  
 DB 308 CSVMHRAHNNHYTKQSLSPGK 330

RESULT 4  
 S31866  
 Ig gamma-1 chain C region - synthetic  
 C:Species: synthetic  
 A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
 C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
 C:Accession: S31866  
 R:Filpula, D.  
 submitted to the EMBL Data Library, February 1993  
 A:Description: Screening method for protein-protein interactions of cloned gene product  
 A:Reference number: S31866  
 A:Accession: S31866  
 A:Molecule type: mRNA  
 A:Residues: 1-255 <FIL>  
 A:Cross-references: EMBL:X70421; NID:G33068; PIDN:CAA49866.1; PID:G33069  
 C:Keywords: immunoglobulin  
 F:1-22/Region: Escherichia coli outer membrane protein A precursor  
 F:23-255/Region: human Ig gamma-1 chain C region

Query Match 41.2%; Score 1197; DB 4; Length 255;  
 Best Local Similarity 96.9%; Pred. No. 8.4e-75;  
 Matches 213; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 319 RPHCTPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 378  
 DB 30 KTHCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 89  
 QY 379 VEVHNATKPREEQYNSTRYVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 438  
 DB 90 VEVHNATKPREEQYNSTRYVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 149  
 QY 439 QPREPOVYTLPPSRREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 498  
 DB 150 QPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 209  
 QY 499 GSFFLYSKLTVDKSRWQQGNVFSQVMHEALHNNHYTKQSLSPGK 544  
 DB 210 GSFFLYSKLTVDKSRWQQGNVFSQVMHEALHNNHYTKQSLSPGK 255

RESULT 5  
 PT0207  
 Ig gamma chain C region - chimpanzee  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
 C:Accession: PT0207  
 R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.  
 Mol. Immunol. 28, 319-322, 1991  
 A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
 A:Reference number: PT0207; MUID:91287716; PMID:2062315  
 A:Accession: PT0207  
 A:Molecule type: mRNA  
 A:Residues: 1-234 <EHR>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:48-117/Domain: immunoglobulin homology <IMW>

Query Match 39.6%; Score 1150; DB 2; Length 234;  
 Best Local Similarity 92.2%; Pred. No. 1.2e-71;  
 Matches 213; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 313 KFTKARP-----HTCPPCPAEEALCAPSVFLPPPKDITLMISRTPEVTCVVVDVSHE 366  
 DB 4 KVDKVEPKSCDTHTCPPCAAEELGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHE 63  
 QY 367 DPEVKFNWYVDGVEVHNAKTKPREEQNTSYRIVSVLTVLHQDLNGLKEYCKKVSNAKLP 426  
 DB 64 DPEVKFNWYVDGVEVHNAKTKPREEQNTSYRIVSVLTVLHQDLNGLKEYCKKVSNAKLP 123  
 QY 427 VPIETKISKAKGQPREPQVITLPPSRDELTQKQVSLTCLVKGYFIPSDIAVEWESNGQPEN 486  
 DB 124 APIETKISKAKGQPREPQVITLPPSRDELTQKQVSLTCLVKGYFIPSDIAVEWESSQPEN 183  
 QY 487 NYKTRPVLDSGSGFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNYHTOKS 537  
 DB 184 NYKTRPVLDSGSGFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNYHTOKS 234

RESULT 6  
 G2HU  
 IG gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA  
 A:Residues: 1-326 <ELL>  
 A:Cross-references: GB:V00554; GB:J00230; NID:G32759; PIDN:CAB58438.1; PID:96066056  
 A:Note: Lys-326 is probably removed posttranslationally  
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
 J. Immunol. 125, 1048-1054, 1980  
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f  
 A:Reference number: A92809; MUID:81007873; PMID:6774012  
 A:Contents: myeloma protein Til  
 A:Accession: A92809  
 A:Molecule type: protein  
 A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
 A:Note: Trp-156 is at or near the complement-binding site  
 R:Connell, G.B.; Parr, D.M.; Hofmann, T.  
 Can. J. Biochem. 57, 758-767, 1979  
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
 A:Reference number: A90752; MUID:80001357; PMID:113060  
 A:Contents: myeloma protein Zie  
 A:Accession: A90752  
 A:Molecule type: protein  
 A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
 A:Note: this sequence has since been revised  
 R:Hofmann, T.; Parr, D.M.  
 Mol. Immunol. 16, 923-925, 1979  
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
 A:Reference number: A93132; MUID:80114419; PMID:118920  
 A:Contents: Zie  
 A:Accession: A93132  
 A:Molecule type: protein  
 A:Residues: 238-275 <HOF>  
 R:Hofmann, T.; Parr, D.M.  
 submitted to the Atlas, March 1980  
 A:Reference number: A94591  
 A:Contents: annotation: Zie, revisions to residues 25, 59, 60, and 264-268  
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
 ned  
 R:Milstein, C.; Frangione, B.  
 Biochem. J. 121, 217-225, 1971  
 A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
 A:Reference number: A90253; MUID:72033500; PMID:4940472  
 A:Contents: annotation: myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.

A:Reference number: A93157; MUID:69064124; PMID:5782707  
 A:Contents: annotation; Sa, disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG2  
 A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (/)  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMI>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental  
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.6%; Score 1149.5; DB 1; Length 326;  
 Best Local Similarity 70.5%; Pred. No. 2.1e-71;  
 Matches 227; Conservative 20; Mismatches 34; Indels 41; Gaps 6;  
 QY 253 KAMTSNVAQVQDNDKFP-----VYKYP-----GKGGPTL----- 283  
 DB 16 RSTSESTAALGCLVKDYFPPBPTVSMNSGALTSVHTFPAVLQSSGLYSLSSVTVFSSN 75  
 QY 284 -GDGDDTLDDYDEETEGDKVRAITVVKPKAPHTCCPPCPAPEALGAPSVFLPPP 342  
 DB 76 FGQTQYTCNVDPKPS-SNTKVDKTVKRCCKVE-----CPPCPAPPVAG-PSVFLPPP 124  
 QY 343 KPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNTSYRIVSV 402  
 DB 125 KPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNTSYRIVSV 184  
 QY 403 LTVLHQDLNGLKEYCKKVSNAKLPVPIETKISKAKGQPREPQVITLPPSRDELTQKQVSL 462  
 DB 185 LTVLHQDLNGLKEYCKKVSNAKLPVPIETKISKAKGQPREPQVITLPPSRDELTQKQVSL 244  
 QY 463 TCLVKGYFIPSDIAVEWESNGQPENNYKTRPVLDSGSGFFLYSKLTVDKSRWQQGNVFNCS 522  
 DB 245 TCLVKGYFIPSDIAVEWESNGQPENNYKTRPVLDSGSGFFLYSKLTVDKSRWQQGNVFNCS 304  
 QY 523 SVMHEALHNYHTOKSLSLSPGK 544  
 DB 305 SVMHEALHNYHTOKSLSLSPGK 326

RESULT 7  
 A23511  
 IG gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:G33070; PIDN:CAA27268.1; PID:G577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 39.3%; Score 1141.5; DB 2; Length 377;  
 Best Local Similarity 63.6%; Pred. No. 9.1e-71;  
 Matches 232; Conservative 19; Mismatches 57; Indels 57; Gaps 8;





```

Db      327 K 327

RESULT 10
IGHWHI
Ig gamma-3 heavy chain disease proteins - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C:Accession: A90442; A92219; A90198; A93915; A02149
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A:Reference number: A90442; MUID:81021548; PMID:6774747
A:Contents: heavy chain disease protein Wis
A:Accession: A90442
A:Molecule type: protein
A:Residues: 1-289 <FRA>
A>Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A>Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A>Note: the sequence of residues 42-76 was taken from the reference that follows
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A:Reference number: A92219; MUID:77118561; PMID:402363
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
A:Accession: A92219
A:Molecule type: protein
A:Residues: 12-97 <MIC>
A>Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
R:Wolfsenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A:Reference number: A90198; MUID:77021516; PMID:823945
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A:Accession: A90198
A:Molecule type: protein
A:Residues: 59-125,'EB',128-226,228-289 <WOL>
A>Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A:Reference number: A93915; MUID:82247835; PMID:6808505
A:Contents: heavy chain disease protein Omm
A:Accession: A93915
A:Molecule type: mRNA
A:Residues: 12-70;72-114;116-125,'E',127-133,'L',135-136,'E',138,'Y',140-154,'D',156-157
A>Note: a carboxyl-terminal Lys is removed posttranslationally
A>Note: this sequence may represent an allelic form or another gamma chain subclass
C:Comment: The heavy chain disease protein Wis is shown.
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglytamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:6,140/Binding site: carbohydrate (Asn) #status experimental

Query Match      38.4%; Score 1115; DB 1; Length 289;
Best Local Similarity 90.6%; Pred. No. 4.le-69;
Matches 203; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY      320 PHTCPCPAPEALGASVFLFPKPKDTLMISRTPEVTCVVDVSHDEPKVKNWYVDGV 379
Db      66 PPPCPRCAPPELLGGFVSFLFPKPKDTLMISRTPEVTCVVDVSHDEPKVKNWYVDGV 125
QY      380 EVNNAKTPREEOYNSTRVSVVLTVLHQDMLNGKEYCKVSNKALPVPIETISKAKQ 439
Db      126 QVHNNAKTPREEOQFNSTRVSVVLTVLHQDMLNGKEYCKVSNKALPVPIETISKAKQ 185

QY      440 PREPOVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDG 499
Db      186 PREPOVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDG 245
QY      500 SFFLYSKLTVDKSRWQOGNPFVSCSVMEALHNHYTQKSLSLSPG 543
Db      246 SFFLYSKLTVDKSRWQOGNPFVSCSVMEALHNHYTQKSLSLSPG 289

RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kackskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126
C:Genetics:
A:Gene: IGG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match      31.6%; Score 918; DB 2; Length 328;
Best Local Similarity 67.5%; Pred. No. 1.6e-55;
Matches 170; Conservative 34; Mismatches 36; Indels 12; Gaps 4;

QY      295 YPEDTEGDKVRAIRTVVVFPTKARPHCTPCCPAPEALGASVFLFPKPKDTLMISRT 354
Db      87 HPATTTIKVDK-----RVGTTKTP-PCPTCPACESPG-PSVFIFFPKPKDTLMISRT 136
QY      355 EVTCVVDVSHDEPKVKNWYVDVSHVNAKTKPREEQYNSTRVSVVLTVLHQDMLNGK 414
Db      137 QVTCVVDVSHDEPKVKNWYVDVSHVNAKTKPREEQYNSTRVSVVLTVLHQDMLNGK 196
QY      415 EYKCKVSNKALPVPIETISKAKQPREQVYTLPPSRBEMTKNQVSLTCLVKGFYPSDI 474
Db      197 EFKCKVNNKLPAPITRIISKAKQPREQVYTLPPSRBEMTKNQVSLTCLVKGFYPSDI 256
QY      475 AVEWESNGQ--PENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNPFVSCSVMEALHNH 532
Db      257 DVWQRNQPEPEGNRYRTTTPQDQVDGTYFLYSKFSVDKASWQGGGIFQCAVMEALHNH 316
QY      533 YTQKSLSLSPGK 544
Db      317 YTQKSLSLSPGK 328

RESULT 12
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kackskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124
C:Genetics:
A:Gene: IGG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

```

Query Match 31.6%; Score 918; DB 2; Length 328;  
 Best Local Similarity 67.5%; Pred. No. 1.6e-55;  
 Matches 170; Conservative 34; Mismatches 36; Indels 12; Gaps 4;

Qy 295 YPEDTEGDKVTRATRVKFPPTKARPHTC--PCPAPALGAPSVFLFPKPKDITLMISRT 354  
 Db HPAATTKVDK-----RVGTTKTP-PCPICPACESPG-PSVFIFPKPKDITLMISRT 136

Qy 355 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLDHQLNGK 414  
 Db QVTCVVVDVSDQENPEVQFQSVYVDGVEVHTAQRKPEQFNSTYRVVSVLPIQHQLNGK 196

Qy 415 EYKCKVSNKALPAPIETKISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDI 474  
 Db EFKCKVNNKDLAPITRIISKAKGTREPQVYTLPPHAEELSRKSVITCLVIGFYPPDI 256

Qy 475 AVESWESNGO--PENNYKTTTPPVLDSDGSPFLYSKLTVDKSRWQQGNVFCSSVMHEALH 532  
 Db DVEWQRNGQPEPEGNRYRTTPQDDVDGTFLYSKFSVDKASWQGGGIFQCAVMHEALH 316

Qy 533 YTKSLSLSPGK 544  
 Db YTKSLSLSPGK 328

RESULT 13  
 GHRB

Ig gamma chain C region - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 24-Apr-1984 #sequence revision 15-Nov-1984 #text change 16-Jul-1999  
 C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
 Immunogenetics 18, 387-397, 1983  
 A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplob  
 A:Reference number: A91749; MUID:84030930; PMID:6313520  
 A:Accession: A91749  
 A:Molecule type: mRNA  
 A:Residues: 1-323 <BER>  
 A:Note: This sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr  
 R:Pratt, D.M.; Mole, L.E.  
 Biochem. J. 151, 337-349, 1975  
 A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglob  
 A:Reference number: A90290; MUID:76135469; PMID:1243651  
 A:Accession: A90290  
 A:Molecule type: protein  
 A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
 A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chain  
 A:Reference number: A93928; MUID:83299917; PMID:6193512  
 A:Accession: A93928  
 A:Molecule type: mRNA  
 A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
 A:Cross-references: GB:M16426; NID:9165111; PIDN:AAA31289.1; PID:G165112  
 A:Note: This sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic mark  
 R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
 Biochem. J. 116, 249-259, 1970  
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin  
 A:Reference number: A90245; MUID:70110015; PMID:5461106  
 A:Accession: A90245  
 A:Molecule type: protein  
 A:Residues: 132-143, 'E', 145-161 <FRU>  
 R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
 in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell.  
 A:Reference number: A94416  
 A:Accession: A94416  
 A:Molecule type: protein  
 A:Residues: 129-131, 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, 'Q',  
 A:Note: This has the e15 allotypic marker, 188-Ala  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-82/Domain: immunoglobulin homology <IM1>  
 F:130-199/Domain: immunoglobulin homology <IM2>  
 F:236-303/Domain: immunoglobulin homology <IM3>  
 F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.4%; Score 913; DB 1; Length 323;  
 Best Local Similarity 66.3%; Pred. No. 3.4e-55;  
 Matches 167; Conservative 31; Mismatches 40; Indels 14; Gaps 2;

Qy 295 YPEDTEGDKVTRATRVKFPPTKARPHTC--PCPAPALGAPSVFLFPKPKDITLMISR 352  
 Db HPAATTKVDKTA-----PSTCKPTCPPELLGGSPVFIFPKPKDITLMISR 131

Qy 353 TPETCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLDHQLN 412  
 Db TPEVTCVVVDVSDQDDPEVQFTYINNEQVTRAPPUREQQFNSTIRVSTLPIQHQLN 191

Qy 413 GKEYCKVSNKALPAPIETKISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS 472  
 Db GKEFKCKVNNKALPAPIETKISKAKGQPLEPKVYTTGPPPREELSSRSVSLTCMNGFYPS 251

Qy 473 DIAVWESNGQPNENYKTTTPPVLDSDGSPFLYSKLTVDKSRWQQGNVFCSSVMHEALH 532  
 Db DISVWEKNGKAEDNYKTTTPAVLDSDGSPFLYNKLSVPTSEWQRGVDFTCSVMHEALH 311

Qy 533 YTKSLSLSPGK 544  
 Db YTKSLSLSPGK 323

RESULT 14  
 I47162

Ig gamma 4 chain constant region - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence revision 21-Feb-1997 #text change 21-Jan-2000  
 C:Accession: I47162  
 R:Kacskovics, I.; Sun, J.; Butler, J.E.  
 J. Immunol. 153, 3565-3573, 1994  
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
 A:Reference number: I47158; MUID:95015845; PMID:7930579  
 A:Accession: I47162  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-277 <KAC>  
 A:Cross-references: EMBL:U03782; NID:G433129; PIDN:AAA52220.1; PID:G433130  
 C:Genetics:  
 A:Gene: IgG4  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 31.3%; Score 911; DB 2; Length 277;  
 Best Local Similarity 67.1%; Pred. No. 3.7e-55;  
 Matches 169; Conservative 33; Mismatches 38; Indels 12; Gaps 4;

Qy 295 YPEDTEGDKVTRATRVKFPPTKARPHTC--PCPAPALGAPSVFLFPKPKDITLMISRT 354  
 Db HPATTTTKVDK-----RVGTTKTP-PCPICPACESPG-PSAFIFPKPKDITLMISRT 85

Qy 355 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLDHQLNGK 414  
 Db KVTCTVVVDVSDQENPEVQFQSVYVDGVEVHTAQRKPEQFNSTYRVVSVLPIQHQLNGK 145

Qy 415 EYKCKVSNKALPAPIETKISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDI 474  
 Db EFKCKVNNKDLAPITRIISKAKGTREPQVYTLPPPTBELSKSVTLTCLVTGFPDDI 205

Qy 475 AVESWESNGO--PENNYKTTTPPVLDSDGSPFLYSKLTVDKSRWQQGNVFCSSVMHEALH 532  
 Db DVEWQRNGQPEPEGNRYRTTPQDDVDGTFLYSKLAVDKASWQRGDTFQCAVMHEALH 265

Qy 533 YTKSLSLSPGK 544  
 Db YTKSLSLSPGK 323

Search completed: October 14, 2003, 06:17:17  
Job time : 27.3023 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 15,1814 Seconds  
(without alignments)  
1685.124 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLLLLLLSPHPHC.....MHEALHHYTKSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1693	58.3	1 GPBA HUMAN	P07359 homo sapien
2	1197.5	41.2	1 GCI_HUMAN	P01857 homo sapien
3	1149.5	39.6	1 GC2_HUMAN	P01859 homo sapien
4	1137	39.1	1 GC4_HUMAN	P01861 homo sapien
5	1120	38.5	1 GC3_HUMAN	P01860 homo sapien
6	913	31.4	1 GC_RABIT	P01870 oryctolagus
7	887	30.5	1 GC2_CAVPO	P01862 cavia porce
8	843	29.0	1 GC3_MOUSE	P22436 mus musculus
9	832	28.6	1 GC3M_MOUSE	P03987 mus musculus
10	831	28.6	1 GCI_RAT	P20759 rattus norv
11	828	28.5	1 GCB_RAT	P20761 rattus norv
12	823	28.3	1 GCC_RAT	P20762 rattus norv
13	820.5	28.2	1 GCI_MOUSE	P01868 mus musculus
14	815.5	28.1	1 GC1M_MOUSE	P01869 mus musculus
15	812	27.9	1 GCA_RAT	P20760 rattus norv
16	805	27.7	1 GCAA_MOUSE	P01863 mus musculus
17	802	27.6	1 GCAB_MOUSE	P01864 mus musculus
18	800	27.5	1 GCAM_MOUSE	P01865 mus musculus
19	767.5	26.4	1 GCB_MOUSE	P01866 mus musculus
20	762.5	26.2	1 GCBM_MOUSE	P01867 mus musculus
21	359.5	12.4	1 EPC_MOUSE	P06336 mus musculus
22	358	12.3	1 MUC_HUMAN	P01871 homo sapien
23	356	12.3	1 MUC2_HUMAN	P04220 homo sapien
24	354	12.2	1 EPC_RAT	P01855 rattus norv
25	353.5	12.2	1 EPC_HUMAN	P01854 homo sapien
26	352	12.1	1 MUC_MOUSE	P01872 mus musculus
27	342	11.8	1 MUCM_MOUSE	P01873 mus musculus
28	338	11.6	1 MUC_RABIT	P03988 oryctolagus
29	331.5	11.4	1 MUC_CANFA	P01874 canis famli
30	329	11.3	1 MUC_MESAU	P06337 mesocricetu
31	328	11.3	1 MUCM_RABIT	P04221 oryctolagus
32	322	11.1	1 MUC_SUNMU	P20768 suncus muri
33	301	10.4	1 A2GL_HUMAN	P02750 homo sapien

34	295	10.2	438	1	HVC2_HETFR	P23085 heterodontu
35	293	10.1	299	1	ALC_RABIT	P01879 oryctolagus
36	289.5	10.0	446	1	MUC_CHICK	P01875 gallus gall
37	286	9.8	353	1	ALCI_HUMAN	P01876 homo sapien
38	283	9.7	438	1	HVC3_HETFR	P23087 heterodontu
39	279	9.6	353	1	ALCI_GORGO	P20758 gorilla gor
40	277.5	9.5	340	1	ALC2_HUMAN	P01877 homo sapien
41	274	9.4	370	1	HVC1_HETFR	P23084 heterodontu
42	274	9.4	461	1	HVC3_HETFR	P23088 heterodontu
43	271	9.3	393	1	HVC3_HETFR	P23086 heterodontu
44	264.5	9.1	646	1	FLRI_HUMAN	Q9nzu1 homo sapien
45	259	8.9	344	1	ALC_MOUSE	P01878 mus musculu

## ALIGNMENTS

RESULT 1	GPBA_HUMAN	STANDARD;	PRT;	626 AA.
AC	P07359; O9HDC7;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, last sequence update)			
DT	15-SEP-2003 (Rel. 42, last annotation update)			
DE	Platelet glycoprotein Ib alpha chain precursor (Glycoprotein Ib alpha)			
DE	(GP-Ib alpha) (GP1BA) (CD42B-alpha) (CD42B) [Contains:			
DE	Glycocalicin].			
GN	GP1BA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1] _TaxID=9606;			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RM	MEDLINE=87289655; PubMed=3303030;			
RA	Lopez J.A., Chung D.W., Fujikawa K., Hagen P.S., Papayannopoulou T.,			
RA	Roth G.J.;			
RT	"Cloning of the alpha chain of human platelet glycoprotein Ib: a			
RT	transmembrane protein with homology to leucine-rich alpha 2-			
RT	glycoprotein.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RM	MEDLINE=89025874; PubMed=2845978;			
RA	Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;			
RT	"Structure of the human blood platelet membrane glycoprotein Ib alpha			
RT	gene.";			
RL	Biochem. Biophys. Res. Commun. 156:389-395(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT PHE-86.			
RM	MEDLINE=22034223; PubMed=12038791;			
RA	Matsubara Y., Murata M., Moriki T., Yokoyama K., Watanabe N.,			
RA	Nakajima H., Handa M., Kawano K., Aoki N., Yoshino H., Ikeda Y.;			
RT	"A novel polymorphism, 70Leu/Phe, disrupts a consensus Leu residue			
RT	within the leucine-rich repeat sequence of platelet glycoprotein			
RT	Ibalpha.";			
RL	Thromb. Haemost. 87:867-872(2002).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT MET-161.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,			
RA	Nickerson D.A.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 17-315.			
RM	MEDLINE=87289654; PubMed=3497398;			
RA	Titani K., Takio K., Handa M., Ruggeri Z.M.;			
RT	"Amino acid sequence of the von Willebrand factor-binding domain of			
RT	platelet membrane glycoprotein Ib.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).			
RN	[6]			
RP	DISULFIDE BONDS.			
RM	MEDLINE=91301149; PubMed=2070794;			
RA	Hess D., Schaller J., Rickli E.B., Clemetson K.J.;			

RT "Identification of the disulphide bonds in human platelet  
glycocalicin."; Eur. J. Biochem. 199:389-393(1991).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
RX MEDLINE=92171415; PubMed=12183630;  
RA Huizinga B.G., Tsuji S., Romijn R.A., Schiphorst M.E., de Groot P.G.,  
SA Sixma J.J., Gros P.;  
RT "Structures of glycoprotein Ibalpha and its complex with von  
Willebrand factor A1 domain."; Science 297:1176-1179(2002).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 17-304, AND SULFATION OF  
TYR-292; TYR-294 AND TYR-295.  
RX MEDLINE=92217968; PubMed=12087105;  
RA Uff S., Clemetson J.M., Harrison T., Clemetson K.J., Ensley J.;  
RT "Crystal structure of the platelet glycoprotein Ibalpha N-terminal  
domain reveals an unmasking mechanism for receptor activation."; J.  
Biol. Chem. 277:35657-35663(2002).  
RN [9]  
RP VARIANT SIBA MET-161.  
RX MEDLINE=92265982; PubMed=1586750;  
RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,  
SA Ruggeri Z.M.;  
RT "Genetic and structural characterization of an amino acid dimorphism  
in glycoprotein Ib alpha involved in platelet transfusion  
refractoriness."; Blood 79:3086-3090(1992).  
RN [10]  
RP VARIANT BSS PHE-73.  
RX MEDLINE=92110577; PubMed=1730088;  
RA Miller J.L., Lyle V.A., Cunningham D.;  
RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein  
Ib alpha leucine tandem repeat occurring in patients with an  
autosomal dominant variant of Bernard-Soulier disease."; Blood  
79:439-446(1992).  
RN [11]  
RP VARIANT BSS VAL-172.  
RX MEDLINE=93388851; PubMed=7690774;  
RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,  
SA de Marco L., Ruggeri Z.M.;  
RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib  
alpha resulting in the Bernard-Soulier syndrome."; J. Clin. Invest.  
92:1213-1220(1993).  
RN [12]  
RP VARIANT BSS SER-225.  
RX MEDLINE=95118882; PubMed=7819107;  
RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,  
SA Ribera A., Gallardo D.;  
RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha  
gene is associated with Bernard-Soulier syndrome."; Br. J. Haematol.  
88:839-844(1994).  
RN [13]  
RP VARIANT PSEUDO-VMD VAL-249.  
RX MEDLINE=91271273; PubMed=2052556;  
RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
RT "Mutation in the gene encoding the alpha chain of platelet  
glycoprotein Ib in platelet-type von Willebrand disease."; Proc. Natl.  
Acad. Sci. U.S.A. 88:4761-4765(1991).  
RN [14]  
RP VARIANT PSEUDO-VMD VAL-249.  
RX MEDLINE=93253059; PubMed=8486780;  
RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
RT "Expression of the phenotypic abnormality of platelet-type von  
Willebrand disease in a recombinant glycoprotein Ib alpha fragment."; J.  
Clin. Invest. 91:2133-2137(1993).  
RN [15]  
RP VARIANT PSEUDO-VMD VAL-255.  
RX MEDLINE=93214031; PubMed=8384898;  
RA Russell S.D., Roth G.J.;  
RT "Pseudo-von Willebrand disease: a mutation in the platelet  
glycoprotein Ib alpha gene associated with a hyperactive surface  
receptor.";

Blood 81:1787-1791(1993).  
RN [16]  
RP VARIANT BSS LEU-195 DEL.  
RX MEDLINE=95178321; PubMed=7873390;  
RA de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
SA Chevalier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
RT "A three-base deletion removing a leucine residue in a leucine-rich  
repeat of platelet glycoprotein Ib alpha associated with a variant of  
Bernard-Soulier syndrome (Nancy I)."; Br. J. Haematol. 89:386-396(1995).  
RN [17]  
RP VARIANT BSS ARG-81.  
RX MEDLINE=98303759; PubMed=9639514;  
RA Kenny D., Jonsson O.G., Mordeck P.A., Montgomery R.R.;  
RT "Naturally occurring mutations in glycoprotein Ibalpha that result in  
defective ligand binding and synthesis of a truncated protein."; Blood  
92:175-183(1998).  
RN [18]  
RP VARIANTS HIS-72 AND MET-161.  
RX MEDLINE=99318093; PubMed=10391209;  
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
SA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nimesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
SA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
of human genes."; Nat. Genet. 22:231-238(1999).  
RN [19]  
RP ERRATUM.  
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
SA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nimesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
SA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
of human genes."; Nat. Genet. 22:231-238(1999).  
RN [20]  
RP VARIANT BSS PRO-145.  
RX MEDLINE=99189763; PubMed=10089893;  
RA Koskela S., Partanen J., Salmi T.T., Kekomaki R.;  
RT "Molecular characterization of two mutations in platelet glycoprotein  
(GP) Ib alpha in two Finnish Bernard-Soulier syndrome families."; Eur. J. Haematol. 62:160-168(1999).  
RN [21]  
RP FUNCTION: GP-Ib, a surface membrane protein of platelets,  
participates in the formation of platelet plugs by binding to the  
A1 domain of von Willebrand factor, which is already bound to the  
subendothelium.  
CC -! SUBUNIT: Heterodimer composed of GP-Ib alpha and beta; disulfide  
linked. GP-IX is complexed with the GP-Ib heterodimer via a non  
covalent linkage.  
CC -! SUBCELLULAR LOCATION: Type I membrane protein.  
CC -! PTM: Glycocalicin, which is approximately coextensive with the  
extracellular part of the molecule, is cleaved off by calpain  
during platelet lysis.  
CC -! POLYMORPHISM: Position 161 is associated with platelet-specific  
allotrans Siba. Siba(-) has Thr-161 and Siba(+) has Met-161.  
CC -! DISEASE: Defects in GPIBA are one of the causes of Bernard-Soulier  
syndrome (BSS). BSS patients have unusually large platelets and  
have a clinical bleeding tendency.  
CC -! DISEASE: Defects in GPIBA are one of the causes of von Willebrand  
disease (VWD) known as platelet-type von Willebrand disease or  
pseudo-von Willebrand disease. This autosomal dominant bleeding  
disorder is caused by an increased affinity of GP-Ib for soluble  
vWF resulting in impaired hemostatic function due to the removal  
of vWF from the circulation.  
CC -! MISCELLANEOUS: Platelet activation apparently involves disruption  
of the macromolecular complex of GP-Ib with the platelet  
glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-  
binding protein.  
CC -! MISCELLANEOUS: Binding sites for von Willebrand factor and  
thrombin (the latter site with unknown function) are in the amino-  
terminal part of the molecule.  
CC -! SIMILARITY: Contains 6 leucine-rich (LRR) repeats.



DR MM; 147100; --  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003623; P:antigen binding activity; TAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig ci.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGC1; 2.  
DR PROSITE; PS00835; IG\_LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
3D-structure.  
FT NON\_TER 1 1  
FT DOMAIN 1 98  
FT HINGE. 99 110  
FT CH2. 111 223  
FT CH3. 224 330  
FT DISULFID 27 83  
FT DISULFID 103 103  
FT DISULFID 109 109  
FT DISULFID 112 112  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT CARBOHYD 180 180  
FT MOD\_RES 330 330  
FT VARIANT 97 97  
FT VARIANT 239 239  
FT VARIANT 241 241  
FT STRAND 122 126  
FT HELIX 130 134  
FT TURN 136 137  
FT STRAND 141 147  
FT STRAND 157 162  
FT TURN 163 164  
FT STRAND 165 166  
FT TURN 168 171  
FT STRAND 176 179  
FT TURN 180 181  
FT STRAND 182 190  
FT HELIX 193 197  
FT TURN 198 199  
FT STRAND 202 207  
FT TURN 209 210  
FT STRAND 215 219  
FT STRAND 227 227  
FT STRAND 230 234  
FT STRAND 238 242  
FT STRAND 245 256  
FT STRAND 260 265  
FT STRAND 270 270  
FT STRAND 274 276  
FT STRAND 280 281  
FT TURN 283 284  
FT STRAND 287 296  
FT STRAND 297 301  
FT TURN 302 303  
FT STRAND 305 312  
FT TURN 313 314  
FT TURN 316 317  
FT STRAND 320 325  
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
Query Match 41.2%; Score 1197.5; DB i; Length 330;  
Best Local Similarity 73.1%; Pred. No. 4.8e-78;  
Matches 236; Conservative 11; Mismatches 37; Indels 39; Gaps 6;  
QY 253 KAMTSNVAIVCDNSDKP-----VYKYPGKCGPTLGDGCTDLYDYYPE 297  
DB 16 KSTSGTAAIGLVKVDYFPFVTVSWNSGALTSVHTFFA-----VLQSSG---LYLSLSSV 68

QY 298 EDTGDKVRAIRTVV-----KPTTKARP-----HTCPPCAPALGAPSVFLFP 341  
DB 69 VTFPSSL-GTQYICNVNHPKSNITKVDKKVBPKSCDKTHTCPCCPAPELLGSPVFLFP 127  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
DB 128 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 187  
QY 402 VLTVLHQLDMLNGKEYCKVSKNKAIPVPIETKISKAKQGPPEQVYITLPPSRREMTKNQVS 461  
DB 188 VLTVLHQLDMLNGKEYCKVSKNKAIPVPIETKISKAKQGPPEQVYITLPPSRDELTKNQS 247  
QY 462 LTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRWQQGNVFS 521  
DB 248 LTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRWQQGNVFS 307  
QY 522 CSVMHEALHNHYTQKSLSLSPGK 544  
DB 308 CSVMHEALHNHYTQKSLSLSPGK 330  
RESULT 3  
GC2 HUMAN  
ID GC2 HUMAN STANDARD; PRT; 326 AA.  
AC P01859;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE IG gamma-2 chain C region.  
GN IGHG2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 2-326 FROM N.A.  
RX MEDLINE=82197621; PubMed=6804948;  
RA Ellison J.W., Hood L.E.;  
RT "Linkage and sequence homology of two human immunoglobulin gamma heavy chain constant region genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
RN [2]  
RP SEQUENCE OF 88-115 FROM N.A.  
RC TISSUE=Fetal liver;  
RX MEDLINE=83001943; PubMed=6811139;  
RA Takahashi N., Ueda S., Obata M., Nikaido T., Nakai S., Honjo T.;  
RT "Structure of human immunoglobulin gamma genes: implications for evolution of a gene family.";  
RL Cell 29:671-679(1982).  
RN [3]  
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
RC TISSUE=Fetal liver;  
RX MEDLINE=84235992; PubMed=6329676;  
RA Krawinkel U., Rabbitts T.H.;  
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass genes.";  
RL EMBO J. 1:403-407(1982).  
RN [4]  
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
RX MEDLINE=81007873; PubMed=6774012;  
RA Wang A.-C., Tung E., Fudenberg H.H.;  
RT "The primary structure of a human IGG2 heavy chain: genetic, evolutionary, and functional implications.";  
RL J. Immunol. 125:1048-1054(1980).  
RN [5]  
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
RX MEDLINE=80001357; PubMed=113050;  
RA Connell G.E., Parr D.M., Hofmann T.;  
RT "The amino acid sequences of the three heavy chain constant region domains of a human IGG2 myeloma protein.";  
RL Can. J. Biochem. 57:758-767(1979).



RP SEQUENCE OF 238-275 (ZIE).  
RX MEDLINE=80114419; PubMed=118920;  
RA Hofmann T., Parr D.M.;  
RT "A note of the amino acid sequence of residues 381-391 of human  
immunoglobulin gamma chains."  
RL Mol. Immunol. 16:923-925(1979).  
RN [7].  
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
RA Hofmann T., Parr D.M.;  
RL Submitted (MAR-1980) to the PIR data bank.  
RN [8].  
RP SEQUENCE OF 1-121 (DOT).  
RX MEDLINE=95255298; PubMed=7737190;  
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
RT "Characterization of the two unique human anti-flavin monoclonal  
immunoglobulins."  
RL Eur. J. Biochem. 228:886-893(1995).  
RN [9].  
RP DISULFIDE BONDS.  
RX MEDLINE=72033500; PubMed=4940472;  
RA Milstein C., Frangione B.;  
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2."  
RL Biochem. J. 121:217-225(1971).  
RN [10].  
RP DISULFIDE BONDS.  
RX MEDLINE=69064124; PubMed=5782707;  
RA Frangione B., Milstein C., Fink J.R.L.;  
RT "Structural studies of immunoglobulin G."  
RL Nature 221:145-148(1969).  
RN [11].  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; J00230; AAB59393.1; -;  
DR PIR; A93906; G2HU;  
DR HSP; P01857; IFCL;  
DR Genew; HGNC:5526; IGHG2.  
DR MIM; 147110; -;  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003823; F:antigen binding activity; TAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig-cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGH1; 2.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON TER 1 109  
FT DOMAIN 1 98 CHI.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 219 CH2.  
FT DOMAIN 220 326 CH3.  
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 83  
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT SITE 156 156  
FT MOD RES 326 326  
FT VARIANT 60 60  
FT CONFLICT 109 109  
AT OR NEAR THE COMPLEMENT-BINDING SITE.  
REMOVED POST-TRANSLATIONALLY (PROBABLE).  
S -> A (IN MYELOMA PROTEINS IIL & ZIE).  
/FTID=VAR\_003889.  
C -> S (IN REF. 3).

SQ SEQUENCE 326 AA; 35084 MW; 8310878C6878CF9C CRC64;  
Query Match 39.6%; Score 1149.5; DB 1; Length 326;  
Best Local Similarity 70.5%; Pred. No. 1.2e-74;  
Matches 227; Conservative 20; Mismatches 34; Indels 41; Gaps 6;  
QY 253 KAMTSNVASVQCDNSDKFP-----YKYP-----GKGCPTL----- 283  
DB 16 RSTSESTAALGLVKDYFPEPTVSVNSGALTSVHTFPAVLQSSGLYSLSVTVTPSN 75  
QY 284 -GDEGTDLDYYPEDTEGDKVTRTRTVKPPHTPCPCPAPALGAPSVFLPPP 342  
DB 76 FGQTQVTCNVDHKP-SNTKVDKTKVERKCCVE-----CPPCPAPPVAG-PSVFLPPP 124  
QY 343 KPQDTLMISRTPEVTCVAVDVSHEDPEVKFNKVDGVEVHNKTPRREOVNSTRVVSU 402  
DB 125 KPQDTLMISRTPEVTCVAVDVSHEDPEVQFNKVDGVEVHNKTPRREOVNSTRVVSU 184  
QY 403 LTVLHQDWLNGKEYKCKVSNKALPVPKEITISKAKGQPREPOVYTLPPSREEMTKNQVSL 462  
DB 185 LTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPOVYTLPPSREEMTKNQVSL 244  
QY 463 TCLVKGFPESDI AVEWESNGQPNNTKTPPVLDSDGSPFLYSKLTVDKSRHQQGVFSC 522  
DB 245 TCLVKGFPESDI AVEWESNGQPNNTKTPPVLDSDGSPFLYSKLTVDKSRHQQGVFSC 304  
QY 523 SVMHEALNHYTKSLSPGK 544  
DB 305 SVMHEALNHYTKSLSPGK 326  
RESULT 4  
GC4\_HUMAN STANDARD; PRT; 327 AA.  
ID GC4\_HUMAN  
AC P01861; 1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ig gamma-4 chain C region.  
GN IGHG4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCB1\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83157104; PubMed=6299662;  
RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene."  
RL DNA 1:11-18(1981).  
RN [2]  
RP SEQUENCE OF 1-30 AND 81-326.  
RX MEDLINE=70207560; PubMed=4192699;  
RA Pink J.R.L., Burtley S.H., de Vries G.M., Milstein C.;  
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
constant region of a gamma 4 chain."  
RL Biochem. J. 117:33-47(1970).  
RN [3]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; K01316; AAB59394.1; ALT\_INIT.  
DR PIR; A90933; G4HU.  
DR PDB; 1ADQ; 16-SEP-98.  
DR Genew; HGNC:5528; IGHG4.  
DR MIM; 147130; -;  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003823; F:antigen binding activity; TAS.

```

DR GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IgG1; 2.
DR PROSITE: PS00835; IG LIKE; 2.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 220
FT DOMAIN 221 327
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 39.1%; Score 1137; DB 1; Length 327;
Best Local Similarity 75.1%; Pred. No. 9.4e-74;
Matches 226; Conservative 17; Mismatches 42; Indels 16; Gaps 5;

QY 246 WKQGVVDVKAMTSNVA5QC--DNSDKFPVVKYFGKGCFTLGDEGDTLDLYDYPEEDTEGD 303
DB 41 WNSG-----ALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTGYTCNVDHKP-SNTKVD 95
QY 304 KVRATRTVTKFPTKARPHCTPCPCAPAPALGAPSVFLFPKPKDTLMISRTPEVTCVVVDV 363
DB 96 K-----RVESYKGP-PCPSCAPEFLGSPVFLFPKPKDTLMISRTPEVTCVVVDV 146
QY 364 SHEDPEVKFMYVDGVEVHNAKTPREQKNSTYRVVSVLTVLHQDMLNKEYKCKVSNK 423
DB 147 SOEDPEVQFMYVDGVEVHNAKTPREQKNSTYRVVSVLTVLHQDMLNKEYKCKVSNK 206
QY 424 ALPVPKEKTSKAGQPREQVYTLPPSRSEMTKNQVSLTCLVKGFYPSDIAVWESNGQ 483
DB 207 GLPSSIETKTSKAGQPREQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVWESNGQ 266
QY 484 PENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPG 543
DB 267 PENNYKTPPVLDSDGSGFFLYSLRTVDKSRWQEGNVFCSCVMHEALHNHYTQKSLSLSPG 326
QY 544 K 544
DB 327 K 327

RESULT 5
GC3_HUMAN
ID GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DD 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RL gamma 3 heavy-chain disease protein Wis.";
RN Biochemistry 19:4304-4308(1980).
RP REVISIONS TO 12-97 (PROTEIN WIS).

RX MEDLINE=77118561; PubMed=402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
RT quadruplication of a 15-amino acid residue basic unit.";
RN J. Biol. Chem. 252:883-889(1977).
RN [3]
RP REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN WIS).
RX MEDLINE=77021516; PubMed=823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the Fc fragment of immunoglobulin G3.";
RN Biochem. Biophys. Res. Commun. 71:907-914(1976).
RN [4]
RP SEQUENCE FROM N.A. (DISEASE PROTEIN WIS).
RX MEDLINE=82447835; PubMed=6808505;
RA Alexander A., Steinmetz M., Barrisault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RT gene deletion model.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -!- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -!- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF. 2.
CC -!- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -!- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
CC GAMMA-3 HEAVY CHAINS.
CC -!- MISCELLANEOUS: DISEASE PROTEIN WIS MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -!- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; J00231; AAA52805.1; ALT_SEQ.
CC HSP; P01857; IFC1.
CC Genew; HGNC:5527; IGHG3.
CC MIM; 147120; -.
CC GO:0005624; C:membrane fraction; NAS.
CC GO:0003823; F:antigen binding activity; TAS.
CC GO:0006955; F:immune response; NAS.
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003597; Ig_c1.
CC InterPro: IPR003006; Ig_MHC.
CC Pfam: PF00047; Ig; 2.
CC SMART: SM00407; IgG1; 1.
CC PROSITE: PS00835; IG LIKE; 2.
CC PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat;
KW Pyridone carboxylic acid.
FT DOMAIN 12 73
FT DOMAIN 74 183
FT DOMAIN 184 289
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27
FT PYRROLIDONE CARBOXYLIC ACID.
FT N-LINKED (GLCNAC...).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).

```



Db 84 HPAINTKVDKTA-----PSTCKPTCPPELLGSPSVFIFPPKPKDTLMISR 131  
 Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVWSVLTVLHQDWLN 412  
 Db 132 TPEVTCVVVDVSDPEVQFTWYINNEQVTRAPPLREQQFNSTIRVWSTLPITHQDWLR 191  
 Qy 413 GKEYCKVSKNALPVPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPS 472  
 Db 192 GKSPKCKVHNAKPAPIEKTISKARGQPLEKPVYTMGPPEELSSRSVSLTCMNGFYPS 251  
 Qy 473 DIAVESNGQPNENYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVFCVSVNHEALHNH 532  
 Db 252 DISVEWKGKGAEDNYKTTTPAVLDSGSEFLYNKLSVPTSEWQSGDVFICSVNHEALHNH 311  
 Qy 533 YTKSKLSLSPGK 544  
 Db 312 YTKSKISRSPOGK 323

RESULT 7  
 GC2\_CAVPO STANDARD; PRT; 329 AA.  
 ID - GC2\_CAVPO AC P01862;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE OF 1-3.  
 RA Trischmann T.M.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [2]  
 RP SEQUENCE OF 4-68.  
 RX MEDLINE=71058471; PubMed=5538606;  
 RA Birhlein B.K., Husain O.Z., Cebr J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the  
 RT half-cystine joining heavy and light chains.";  
 RN Biochemistry 10:18-25(1971).  
 RN [3]  
 RP SEQUENCE OF 69-133 AND 312-329.  
 RX MEDLINE=71058486; PubMed=5538616;  
 RA Turner K.J., Cebr J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal  
 RT and hinge region cyanogen bromide fragments.";  
 RN Biochemistry 10:9-17(1971).  
 RN [4]  
 RP SEQUENCE OF 134-226.  
 RX MEDLINE=75036072; PubMed=4429665;  
 RA Tracey D.E., Cebr J.J.;  
 RT "Primary structure of the CH2 homology region from guinea pig IgG2  
 RT antibodies.";  
 RN Biochemistry 13:4796-4803(1974).  
 RN [5]  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; PubMed=4609467;  
 RA Trischmann T.M., Cebr J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig IgG2  
 RT antibodies.";  
 RN Biochemistry 13:4804-4811(1974).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71058474; PubMed=4922544;  
 RA Oliveira B., Lamm M.B.;  
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";  
 RN Biochemistry 10:26-31(1971).  
 CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN  
 CC 13 INBRED GUINEA PIGS.

DR PIR: A94553; G2GP.  
 DR HSSP: P01842; 7FAB.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003597; Ig\_c1.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART: SM00407; Igcl; 2.  
 DR PROSITE: PS00835; IG\_LIKE; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON TER 1 1  
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 28 79  
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 142 202  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).  
 FT DISULFID 248 308  
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;  
 Query Match 30.5%; Score 887; DB 1; Length 329;  
 Best Local Similarity 65.9%; Pred. No. 5.8e-56;  
 Matches 166; Conservative 29; Mismatches 49; Indels 8; Gaps 3;  
 Qy 295 YPEEDTEGDK-VRAIRTRVWKFPPTKARPHCTPCPAPEALGAPSVFLPPKPKDTLMISR 353  
 Db 83 HPASSTKVDKTVETPIRT-----PZPBCTCPKPPENLGSPSVFIFPPKPKDTLMISLT 137  
 Qy 354 PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVWSVLTVLHQDWLN 413  
 Db 138 PRVTCVVVDVSDPEVQFTWYINNEQVTRAPPLREQQFNSTIRVWSTLPITHQDWLRG 197  
 Qy 414 KEYCKVSKNALPVPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSD 473  
 Db 198 KEFKCKVYNKALPAPIEKTISKAKGAPMPDVFYTLPPSRDELKSKSVSVTCLINFPAD 257  
 Qy 474 IAVESNGQPN - ENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVFCVSVNHEALHN 531  
 Db 258 IAVESNGQPN - ENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVFCVSVNHEALHN 317  
 Qy 532 HYTKSKLSLSPG 543  
 Db 318 HVTQKISRSPG 329

RESULT 8  
 GC3\_MOUSE STANDARD; PRT; 329 AA.  
 ID - GC3\_MOUSE AC P22436;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-3 chain C region, secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85027161; PubMed=6092053;  
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
 RA Tucker P.M., Blattner F.R.;  
 RT "Structural analysis of the murine IgG3 constant region gene.";  
 RN EMBO J. 3:2041-2046(1984).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; J00451; -; NOT_ANNOTATED_CDS.
CC PIR; B02156; G3MSC.
CC HSSP; P01857; 1FC1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IG1; 2.
CC PROSITE; PS00835; IG LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362
FT DOMAIN 333 338
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 29.0%; Score 843; DB 1; Length 329;
Best Local Similarity 65.5%; Pred. No. 7.8e-53;
Matches 152; Conservative 32; Mismatches 46; Indels 2; Gaps 1;

QY 315 PTKARHTCP--PCPAPEALGAPSVLFPKPKDLMISRTPEVTCVVDVSHEDPEVKF 372
DB 98 PRIPKSTPPGSSCPGNTILGGPSVFIFPPKPKDALMSLTPEKVTGVVDVSEDDPDVHV 157

QY 373 NWYDGVGVHNAKTPREQYNSTYRVWSVLTVLHODWLNGLKYEKCKVSKNKPVPPIET 432
DB 158 SWFVDNKEVHTAWTPREAQYNSTFRVVSALPIQHODWNRGKFKCKVNNKALPAPIET 217

QY 433 ISKAGQPREPQVYTLPPGREMTKQVSLTCLVKGFPSPDIKAVESNGQPPENNYKTP 492
DB 218 ISKPKGRAQTPQVYTLPPGREMTKQVSLTCLVNTFFSEALSVENRNGELEQDYKNT 277

QY 493 PVLDSDGSPFLYSKLTVDKSRWQGVNFCVSVHNEALHNYTKQSLSPGK 544
DB 278 PILDSDGTFLYSKLTVDTSMLQGEIFTCSVVHNEALHNYTKQSLSPGK 329

RESULT 9
GC3M_MOUSE STANDARD; PRT; 398 AA.
ID GC3M_MOUSE
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=8401483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00451; AAB59655.1; -.
CC EMBL; V01526; CAA24767.1; ALT_SEQ.
CC PIR; A02156; G3MSM.
CC HSSP; P01857; 1FC1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IG1; 2.
CC PROSITE; PS00835; IG LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362
FT DOMAIN 333 338
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7264B50A41B95 CRC64;

Query Match 28.6%; Score 832; DB 1; Length 398;
Best Local Similarity 65.2%; Pred. No. 6e-52;
Matches 150; Conservative 32; Mismatches 46; Indels 2; Gaps 1;

QY 315 PTKARHTCP--PCPAPEALGAPSVLFPKPKDLMISRTPEVTCVVDVSHEDPEVKF 372
DB 98 PRIPKSTPPGSSCPGNTILGGPSVFIFPPKPKDALMSLTPEKVTGVVDVSEDDPDVHV 157

QY 373 NWYDGVGVHNAKTPREQYNSTYRVWSVLTVLHODWLNGLKYEKCKVSKNKPVPPIET 432
DB 158 SWFVDNKEVHTAWTPREAQYNSTFRVVSALPIQHODWNRGKFKCKVNNKALPAPIET 217

QY 433 ISKAGQPREPQVYTLPPGREMTKQVSLTCLVKGFPSPDIKAVESNGQPPENNYKTP 492
DB 218 ISKPKGRAQTPQVYTLPPGREMTKQVSLTCLVNTFFSEALSVENRNGELEQDYKNT 277

QY 493 PVLDSDGSPFLYSKLTVDKSRWQGVNFCVSVHNEALHNYTKQSLSP 542
DB 278 PILDSDGTFLYSKLTVDTSMLQGEIFTCSVVHNEALHNYTKQSLSP 327

RESULT 10
GC1_RAT STANDARD; PRT; 326 AA.
ID GC1_RAT
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.

```

```

DR Pfam: PF00047; ig; 3.
DR SMART: SM00407; Igel; 2.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON TER 1 97
FT DOMAIN 1 97
FT DOMAIN 98 112
FT DOMAIN 113 219
FT DOMAIN 220 326
FT DOMAIN 27 82
FT DISULFID 102 102
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 109 109
FT DISULFID 111 111
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 28.6%; Score 831; DB 1; Length 326;
Best Local Similarity 44.8%; Pred. No. 5.5e-52;
Matches 165; Conservative 50; Mismatches 77; Indels 76; Gaps 6;

QY 195 LOENSLYT---IPKGF-----GSHLLPFAFLHGNPWLNCCEILYFRWL 236
DB 17 LKSNMVTGLCLVKGYPPEPTVWNSGALSSGVHTFAVLQSG----- 60

QY 237 QDNAENVYVWQGVVDVAMTSNVASVQCDNSDKFPVKYKPGKPTLGDGDTLDLYDYP 296
DB 61 -----LYLTSSVTFSSTWPSQTVTCNVA-----HP 87

QY 297 EEDTEGDKVTRATRVVFPPTKARPHTCPPCAPALGAPSVLFPKPKDTLMISRTPEV 356
DB 88 ASSTKVDK-----KIVPRNCQDCKPCICTGS-EVSSVFIPPPKPKDVLITLTPKV 138

QY 357 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEY 416
DB 139 TCVVVDISQDDPEVHFNSFVDDVEVHTAQTAPPEQFNSTFRSVELPILHQDWLNGRTF 198

QY 417 KCKVSNKALPVIETISKAKGQPREPOVYTLPPSRREMTKNQVSLCCLVKGYFSPDIAV 476
DB 199 RCKVTSAAFPSPIEKTIKSGPEGTRQVPHVYVMTSPKEEMTQNEVSIITCMVKGFYPPDIY 258

QY 477 EWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQK 536
DB 259 EQWMQGQPENNYKNTPTMTDIDGSIYFLYSKUNVKRKKMQQNTFTCSVLHEGLNHRHTEK 318

QY 537 SLSLSPGK 544
DB 319 SLSHSPGK 326

```

## RESULT 11

```

ID GCB RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC !- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR; PS0018;

```

```

DR HSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; Igel; 2.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON TER 1 96
FT DOMAIN 1 96
FT DOMAIN 124 223
FT DOMAIN 232 328
FT DISULFID 15 15
FT DISULFID 27 80
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 112 112
FT DISULFID 115 115
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 28.5%; Score 828; DB 1; Length 333;
Best Local Similarity 50.1%; Pred. No. 9.2e-52;
Matches 169; Conservative 42; Mismatches 92; Indels 34; Gaps 6;

QY 237 QDNAENVYVWQGVVDVAMTSNVASVQCDNSDKFP-----VYKPG---K 278
DB 2 QTTAPVPLAPGCG---DTTSTVTGLVKGYPPEPTVWNSGALSSDVHTFPVAVLQS 59

QY 279 GCPTLGDGDTLDYD-----YYPEEDTEGDKVTRATRV---VVKPTKARPHTCPCP 327
DB 60 GLYLTSSVTFSSTWPSQTVTCNVAHPASSTVKVKKVERNGGIGHKCPTCP---TCHKCP 116

QY 328 APEALGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTX 387
DB 117 VPGLGGSPVFIIPKPKDILLISQNAKVTCCVVVDVSEEDPDQFSWFNANVEVHTAQTQ 176

QY 388 PREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVIETISKAKGQPREPOVYT 447
DB 177 PREQYNSTFRVSALPTIQHDMWSGKPKCKVNNKALPSPIEKTIKSKGLVKRPQVYV 236

QY 448 LPPSREEMTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKL 507
DB 237 MCPPEQTEQTEVTSVLTCLTSGLFNDIGVETNSGHIKKNYKNTPEWMDSDGSFFMYSKL 296

QY 508 TVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPGK 544
DB 297 NVESRMRSDRAPFCVSVVHGLNHNHVEKSISSRPFGK 333

```

## RESULT 12

```

GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M.; Delmastro-Galfire P.; Waldmann H.; Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2C heavy chain constant
region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X07189; CAA30169.1; -  
 DR HSP: S00847; S00847.  
 DR HSP: P01842; 7PAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00407; IGc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Immunoglobulin domain; Immunoglobulin C region.

FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 222 CH2.  
 FT DOMAIN 223 329 CH3.  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 249 307  
 SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 28.3%; Score 823; DB 1; Length 329;  
 Best Local Similarity 62.0%; Pred. No. 2.1e-51;  
 Matches 150; Conservative 42; Mismatches 46; Indels 4; Gaps 2;

QY 307 ATRT-VKFTKARPTCP---CPAPEALGAPSVFLPFPKDTLMISRTSEVTCVVVD 362  
 D 88 ATKSNLKKIEPRPRPTDTCSDNLRGSPVFPFPKDTLMISRTSEVTCVVVD 147  
 QY 363 VSHEDPEVFNWVVDGVEVHNATKPREQYNSTYRWVSLTVLHODWLNKGEYCKVSN 422  
 D 148 VSEEDPQVSWFVNVVFTATQPHBQLNGTFPRVSTLHIOQDMWMSGKEFKCKVNN 207  
 QY 423 KALPVPKTIKAKQPREPQVYTLPPSRENTKQVSLTCLVKGFPYSDIAVWESNG 482  
 D 208 KDLPSPIKTIKPRGKAR-PQVYTLPPSRENTKQVSLTCLVKGFPYSDIAVWESNG 267  
 QY 483 OPENNYKTPPVLDSDGSFLYSKLTVDKSRWQQGNVFCSSVWHEALHNYTKSLSP 542  
 D 268 ELEQDYNTLPVLDSDGSFLYSKLTVDKSRWQQGNVFCSSVWHEALHNYTKSLSP 327  
 QY 543 GK 544  
 D 328 GK 329

RESULT 13  
 ID GC1\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01568;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-1 chain C region secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;

"Cloning and complete nucleotide sequence of mouse immunoglobulin  
 gamma 1 chain gene.";  
 Cell 18:559-568(1979).

[2]  
 RN SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=800202559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 cloned in a bacterial plasmid.";  
 Gene 9:87-97(1980).

[3]  
 RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salsner W.;  
 "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 heavy chain.";  
 Nucleic Acids Res. 6:3305-3321(1979).

[4]  
 RN SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Adetugbo K.;  
 "Evolution of immunoglobulin subclasses. Primary structure of a  
 murine myeloma gamma1 chain.";  
 J. Biol. Chem. 253:6068-6075(1978).

[5]  
 RN DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
 Biochem. J. 126:837-850(1972).

CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=IP01868-1; Sequence=Displayed;  
 CC Note=May be the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=IP01869-1; Sequence=External;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; V00793; CAA24172.1; -  
 DR EMBL; V00793; CAA24173.1; -  
 DR EMBL; V00793; CAA24174.1; -  
 DR EMBL; V00793; CAA24175.1; -  
 DR EMBL; V00795; CAA24176.1; -  
 DR PIR; A02159; GIMS.  
 DR PDB; 1IGC; 03-JUN-95.  
 DR GlycoSuiteDB; P01868; -  
 DR MGD; MGI:96446; Igh-4.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; IGc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing; 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 127 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).

```
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 /FTIG=CAR_000055.
FT DISULFID 244 302 REMOVED POST-TRANSLATIONALLY.
FT MOD RES 324 324 N -> D (IN REF. 3).
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 28.2%; Score 820.5; DB 1; Length 324;
Best Local Similarity 57.9%; Pred. No. 3e-51;
Matches 146; Conservative 43; Mismatches 48; Indels 15; Gaps 3;

Qy 295 YPEEDTGDGKVRATRVVVKFPTKARPHTCPP--CPAPEALGAPSVFLPPPKDPTLMISR 352
Db 86 HPASSTKVDKKIVPRDC-----GCKPCICTVPE---VSSVFIFPPKPKDVLITL 132
Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREOYNSTIRYVSVLTVLHODWLN 412
Db 133 TPKVTCVVVDISKDPEVQFSWFVDVDEVHTAQTQPREEQFNSTFRSVSELPIMHQDWLN 192
Qy 413 GREYKCKYSNKALPVPPIETKISKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPS 472
Db 193 GREYKCRVNSAFAPIETKISKGRKAPQVYIIPPKQMAKDKVSLCKMTDFPPE 252
Qy 473 DIAVEVNGSQPNKYKTPPVLSDSGFFLSKLTVDKSRWQQNVFSCSVHMEALNHH 532
Db 253 DITVEWQNGQPAENYKNYKQPIWNTNGSYFYVSKLVNOKSNWEAGNTFTCSVLHGLNHS 312
Qy 533 YTKSLSLSPGK 544
Db 313 HTEKSLSHSPGK 324

RESULT 14
GCLIM_MOUSE
ID GCLIM_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
```

```
[4]
SEQUENCE OF 1-44 FROM N.A.
MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
RL -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Membrane-bound;
IsoId=P01869-1; Sequence=Displayed;
Name=Secreted;
IsoId=P01868-1; Sequence=External;
Note=May be the major isoform;
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; V00793; CAA24172.1; -
EMBL; V00793; CAA24173.1; -
EMBL; V00793; CAA24174.1; -
PIR; B02159; GMSM
PDB; 15C8; 23-MAR-99.
PDB; 1AB6; 18-MAR-98.
PDB; 1CL7; 12-JAN-00.
PDB; 1F11; 06-FEB-01.
PDB; 1F58; 29-DEC-99.
PDB; 1KCS; 24-JUL-02.
PDB; 1KCR; 11-MAY-02.
PDB; 25C8; 09-JUL-99.
MGD; MGI:96446; Igh-4.
InterPro; IPR007110; Ig-like.
InterPro; IPR003597; Ig cl.
InterPro; IPR003006; Ig_MHC.
Pfam; PF00047; Ig; 3.
SMART; SM00407; IGcl; 2.
PROSITE; PS00835; IG_LIKE; 3.
PROSITE; PS00290; IG_MHC; 1.
Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
Alternative splicing; Transmembrane; 3D-structure.
NON_TER 1
DOMAIN 1 97 CHI.
DOMAIN 98 110 HINGE.
DOMAIN 111 217 CH2.
DOMAIN 218 324 CH3.
DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 138 198 N-LINKED (GLCNAC. . .).
CARBOHYD 174 174
DISULFID 244 302 POTENTIAL.
DISULFID 340 357 CYTOPLASMIC (POTENTIAL).
DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 28.1%; Score 815.5; DB 1; Length 393;
Best Local Similarity 57.8%; Pred. No. 8.9e-51;
Matches 145; Conservative 43; Mismatches 48; Indels 15; Gaps 3;

Qy 295 YPEEDTGDGKVRATRVVVKFPTKARPHTCPP--CPAPEALGAPSVFLPPPKDPTLMISR 352
Db 86 HPASSTKVDKKIVPRDC-----GCKPCICTVPE---VSSVFIFPPKPKDVLITL 132
Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREOYNSTIRYVSVLTVLHODWLN 412
Db 133 TPKVTCVVVDISKDPEVQFSWFVDVDEVHTAQTQPREEQFNSTFRSVSELPIMHQDWLN 192
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 57.1833 Seconds  
(without alignments)  
2454.925 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLLLLLLLPLPHPHPIC.....MHEALNHYTKQSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp.archaea:\*

2: sp.bacteria:\*

3: sp.fungi:\*

4: sp.human:\*

5: sp.invertebrate:\*

6: sp.mammal:\*

7: sp.mhc:\*

8: sp.organelle:\*

9: sp.phage:\*

10: sp.plant:\*

11: sp.podent:\*

12: sp.virus:\*

13: sp.vertibrate:\*

14: sp.unclassified:\*

15: sp.xvirus:\*

16: sp.bacteriap:\*

17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1693	58.3	626	4	Q8N1F3 homo sapien
2	1687	58.1	626	4	Q8NG39
3	1224.5	42.1	701	4	Q96PQ8
4	1207.5	41.6	471	4	Q8TC77
5	1137.5	39.1	473	4	Q8TC63
6	1137.5	39.1	521	4	Q8N4Y9
7	1126.5	38.8	509	4	Q8N1F7
8	1076.5	37.0	734	11	Q35930
9	970	33.4	677	6	Q28256
10	891	30.7	337	6	Q95W34
11	838	28.8	469	11	Q8R3V9
12	835	28.7	463	11	Q93LC4
13	831	28.6	437	11	Q9R1A4
14	808	27.8	473	11	Q908L4
15	801	27.6	468	11	Q99L31
16	801	27.6	473	11	Q99L25

17	775	26.7	473	11	Q91Z05	Q91Z05 mus musculu
18	773	26.6	474	11	Q8R3H6	Q8R3H6 mus musculu
19	358	12.3	375	4	Q9BSZ1	Q9BSZ1 homo sapien
20	358	12.3	597	4	Q9BU10	Q9BU10 homo sapien
21	358	12.3	597	4	Q9BQB8	Q9BQB8 homo sapien
22	358	12.3	597	4	Q9BBB9	Q9BBB9 homo sapien
23	348	12.0	588	4	Q8WUX4	Q8WUX4 homo sapien
24	348	12.0	613	4	Q96EY0	Q96EY0 homo sapien
25	348	12.0	613	4	Q8WUK1	Q8WUK1 homo sapien
26	348	12.0	614	4	Q96GAG	Q96GAG homo sapien
27	348	12.0	618	4	Q96AA6	Q96AA6 homo sapien
28	342	11.8	613	11	Q8VCX7	Q8VCX7 mus musculu
29	302.5	10.4	342	11	Q91XL1	Q91XL1 mus musculu
30	301	10.4	347	4	Q8N4F5	Q8N4F5 homo sapien
31	293.5	10.1	1328	5	Q21043	Q21043 caenorhabdi
32	287	9.9	486	11	Q91Z07	Q91Z07 mus musculu
33	286	9.8	384	4	Q9UF60	Q9UF60 homo sapien
34	286	9.8	493	4	Q8NCL6	Q8NCL6 homo sapien
35	286	9.8	494	4	Q96K68	Q96K68 homo sapien
36	286	9.8	496	4	Q96KX8	Q96KX8 homo sapien
37	284	9.8	499	4	Q8N5K4	Q8N5K4 homo sapien
38	283	9.7	496	4	Q96DK0	Q96DK0 homo sapien
39	283	9.7	500	4	Q9BRV0	Q9BRV0 homo sapien
40	281	9.7	426	11	Q9DCD9	Q9DCD9 mus musculu
41	281	9.7	497	4	Q8WY24	Q8WY24 homo sapien
42	280	9.6	487	11	Q99KA4	Q99KA4 mus musculu
43	279	9.6	416	4	Q9NPP6	Q9NPP6 homo sapien
44	279	9.6	1527	5	Q9VZ24	Q9VZ24 drosophila
45	277.5	9.5	479	11	Q91WP5	Q91WP5 mus musculu

ALIGNMENTS

RESULT 1

Q8N1F3 ID Q8N1F3 PRELIMINARY; PRT; 626 AA.

AC Q8N1F3;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Glycoprotein Ib (Platelet), alpha polypeptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC027955; AAH27955.1; -.

DR InterPro; IPR001611; LRR\_Cterm.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR000372; LRR\_Nterm.

DR InterPro; IPR003591; LRR\_Typ.

DR InterPro; IPR002965; P-rich\_extensn.

DR Pfam; PF00560; LRR; 1\_

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01462; LRRNT; 1.

```

Db 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNMLTALPPDLKDTTILHLSNLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKQVDTGLPVLGTDLDSHNLQSLPLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKQVDTGLPVLGTDLDSHNLQSLPLGQTLPALTV 120
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLEYTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLEYTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
QY 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKPGKGTGLGDEGDTLDYDYPEEDT 300
Db 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKPGKGTGLGDEGDTLDYDYPEEDT 300
QY 301 EGDKVRATRTVVKPTKARHTCP-----PCPA---PEALGAPSVFLFPPK 343
Db 301 EGDKVRATRTVVKPTKARHTCP-----PCPA---PEALGAPSVFLFPPK 343
QY 344 --PKDTL-----MISRTPEVT 357
Db 359 WTPNFTLHMSITFSKTPKST 379

RESULT 2
Q8NG39 PRELIMINARY; PRT; 626 AA.
AC Q8NG39;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Platelet glycoprotein Ib alpha.
GN GPIBA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsumura Y., Murata M., Sugita K., Ikeda Y.;
RT "Identification of a novel point mutation in platelet glycoprotein
RT Iba, Gly co Ser at residue 233, in a Japanese family with platelet-
RT type von Willebrand disease.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086948; BAC10305.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR002965; P_Rich_extensn.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 5.
DR PROSITE; PS50506; LRR_TYPICAL; 1.
SQ SEQUENCE 626 AA; 68985 MW; C7931FD07458B17P CRC64;

Query Match 58.1%; Score 1687; DB 4; Length 626;
Best Local Similarity 87.1%; Pred. No. 3.4e-131;
Matches 332; Conservative 4; Mismatches 19; Indels 26; Gaps 5;

QY 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNMLTALPPDLKDTTILHLSNLLY 60
Db 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNMLTALPPDLKDTTILHLSNLLY 60

```

```

QY 61 TFSLATLMPYTRLTQNLDRCELTKQVDTGLPVLGTDLDSHNLQSLPLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKQVDTGLPVLGTDLDSHNLQSLPLGQTLPALTV 120
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLEYTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLEYTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
QY 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKPGKGTGLGDEGDTLDYDYPEEDT 300
Db 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKPGKGTGLGDEGDTLDYDYPEEDT 300
QY 301 EGDKVRATRTVVKPTKARHTCP-----PCPA---PEALGAPSVFLFPPK 343
Db 301 EGDKVRATRTVVKPTKARHTCP-----PCPA---PEALGAPSVFLFPPK 343
QY 344 --PKDTL-----MISRTPEVT 357
Db 359 WTPNFTLHMSITFSKTPKST 379

RESULT 3
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Factor VII active site mutant immunonjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF272774; AKS8686.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00047; ig; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00407; IGc1; 1.
DR SMART; SM00020; Tryp_Sfc; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.

```

```

DR PROSITE; PS01187; EGF CA; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 701 AA; 71826 MW; 94AC6CEB42CC992F CRC64;

Query Match 42.1%; Score 1224.5; DB 4; Length 701;
Best Local Similarity 83.3%; Pred. No. 9.1e-93;
Matches 235; Conservative 5; Mismatches 23; Indels 19; Gaps 3;

QY 277 GKGCPTLGDGDTLDYDYPD-----EDTEGDKVRAIRTVVKPTKARP-----HT 322
DB 425 GQGCATVGHGFYTRVSQYIEWLQKMRSEFRPGVLLRA-----PPFGSAEPKSCDKTHT 479

QY 323 CPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 382
DB 480 CPPCPAPELGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 539

QY 383 NAKTKPREQNSTYRVVSVLTVLHQDLNKGKYCKKVSNAKALPVPPIKTIKAKQPRE 442
DB 540 NAKTKPREQNSTYRVVSVLTVLHQDLNKGKYCKKVSNAKALPAPIKTIKAKQPRE 599

QY 443 PQVTLTPSREEMTKNQVSLTCLVKGYFPSDIAVEWESNGQPENNYKTPPPVLDSDGSF 502
DB 600 PQVTLTPSRDELTKNQVSLTCLVKGYFPSDIAVEWESNGQPENNYKTPPPVLDSDGSF 659

QY 503 LYSKLTVDKSRWQQGNVFCSSVMHEALHNHYTQKSLSLSPGK 544
DB 660 LYSKLTVDKSRWQQGNVFCSSVMHEALHNHYTQKSLSLSPGK 701

RESULT 4
Q8TC77
ID Q8TC77 PRELIMINARY; PRT; 471 AA.
AC Q8TC77;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024289; AAH24289.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 41.6%; Score 1207.5; DB 4; Length 471;
Best Local Similarity 64.7%; Pred. No. 1.3e-91;
Matches 244; Conservative 19; Mismatches 43; Indels 71; Gaps 6;

QY 235 WLQDN---AENVYWKQGVQVYKAMTS-----NVASVQCDNSDKFPV---YKY 275
DB 99 YLQMSLRAEDTAVYICARDLRQLTSYWFYDLWGRGLTVTVSSASTKGPSVFFLAPSSKS 158

QY 276 PGKCPITLGDGDDTLDYDYPEDT-----EGDKVRAIRTVVKPT 316

```

```

DB 159 TSGCTAALG---CLVKDYFPEPVTVMNSGALTSVGHTPPAVLQSSGLYSLSVVTVP 214
QY 317 KA-----RPHTPCPAPEALGAPSVFLPPPKPKOT 347
DB 215 SSLGTQTYICNVNHPKSNTRKVDKKVBPKSCDKTHTCPCPAPELGGPSVFLPPPKPKOT 274
QY 348 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 407
DB 275 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 334
QY 408 QDWLNGKEYKCKVSNKALPAPIEKTIKAKQPREPOVYTLPPSRDEMTKNQVSLTCLVK 467
DB 335 QDWLNGKEYKCKVSNKALPAPIEKTIKAKQPREPOVYTLPPSRDELTKNQVSLTCLVK 394
QY 468 GFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVMHE 527
DB 395 GFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVMHE 454
QY 528 ALHNHYTQKSLSLSPGK 544
DB 455 ALHNHYTQKSLSLSPGK 471

RESULT 5
Q8TC63
ID Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025985; AAH25985.1; -.
DR InterPro; IPR000923; BlueCu_1.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 39.1%; Score 1137.5; DB 4; Length 473;
Best Local Similarity 75.1%; Pred. No. 8.6e-86;
Matches 226; Conservative 17; Mismatches 42; Indels 16; Gaps 5;

QY 246 WKQGVYKAMTSNVAQVC--DNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGD 303
DB 187 WNSG----ALTSGVHTPPAVLQSSGLYSLSVVTVPSSSLGTITYTCNVDHKP-SNTKVD 241
QY 304 KVRATRTVVKFPTKARPHTPCPAPEALGAPSVFLPPPKPKOTLMISRTPEVTCVVVDV 363
DB 242 K-----RVESKYGP-PCPSCPAPFLGGPSVFLPPPKPKOTLMISRTPEVTCVVVDV 292
QY 364 SHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKYCKKVSNAK 423
DB 293 SOEQVQFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKYCKKVSNAK 352
QY 424 ALPAPIEKTIKAKQPREPOVYTLPPSRDEMTKNQVSLTCLVKGYFPSDIAVEWESNGQ 483
DB 353 GLPSIETIKAKQPREPOVYTLPPSRDEMTKNQVSLTCLVKGYFPSDIAVEWESNGQ 412

```

[illegible]

RESULT 6	Q8N4Y9	PRELIMINARY;	PRT;	521 AA.
ID	Q8N4Y9			
AC	Q8N4Y9;			
DT	01-OCT-2002 (trEMBLrel. 22, Created)			
DT	01-OCT-2002 (trEMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (trEMBLrel. 23, Last annotation update)			
DE	Hypothetical protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			

RP SEQUENCE FROM N.A.  
RC TISSUE=Primary B-Cells from Tonsils;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR	EMBL; BC033178; AAH33178.1; -	
DR	InterPro; IPR007110; IG-like.	
DR	InterPro; IPR003597; IG_cl.	
DR	InterPro; IPR003006; IG_MHC.	
DR	InterPro; IPR003596; IG_v.	
DR	Pfam; PF00047; IG; 4.	
DR	SMART; SM00407; IGcl; 3.	
DR	SMART; SM00406; IGv; 1	
DR	PROSITE; PS0835; IG_LIKE; 4.	
DR	PROSITE; PS00290; IG_MHC; 2.	
KW	Hypothetical protein.	
SQ	SEQUENCE 521 AA; 57156 MW; 2AC7D22E72D6CAA2 CRC64;	

Query Match	Score	DB 4	Length
Best Local Similarity	63.3%		521
Matches 231; Conservative		20; Mismatches	57; Gaps
			8;

Qy	194	LLOENSLY-----TIPKGFSSHLHPAFUHGHPWLNCEILYFRRWLOQNNAEVYWK	247
Db	200	VLOSSGLYSLSVVTVPPSSLT-----QTYTCN-----VNHKPSNTKVDK	240
Qy	248	QGYDVKAMTSNVASV-----QCDNSDKPFVYKYPKGCGPTLGDGEGDLDLYVYPEED	299
Db	241	R-VELKTPUGDTHHTCPRCPEKSCDTPPPCPRCPEP-KSCCT-----PPPC	285

Qy	300	TEGDKVTRATRTVVKGFTTKARPHUTCPCPAPALGAPSVFLFPPKPKDTLMISRTPEVTCV	359
Db	286	PRCPKPKCDT-----PPCPKCPAPELLLGGPSVFLFPPKPKDTLMISRTPEVTCV	336
Qy	360	VDPVSHEDPEVKFNWYVDGVEVHNAKTKPRBEQYNSTYRVVSVLTVLDHQLMKGKSYKCK	419
Db	337	VDPVSHEDPEVQPKWYVDGVEVHNAKTKPRBEQYNSTYRVVSVLTVLDHQLMKGKSYKCK	396

Qy	420	VSNKALPVIPIEKTISKAKQPREPOVYTLPPSREMTKNQVSLCLVKGYPSDIAVEWE	479
Db	397	VSNKALPAPIEKTISKTKQPREPOVYTLPPSREMTKNQVSLCLVKGYPSDIAVEWE	456
Qy	480	SNQGPENNYKTPPVVLDSGSEFFLYSKLTIVKSRMQQNVFSCSVMHEALHNHYTKQSL	539
Db	457	SSQGPENNYNTIPMLDSDGSEFFLYSKLTIVKSRWQONIFSCSVMHEALHNFTQKSL	516

Qy	540	LSPGK	544
Db	517	LSPGK	521

## RESULT 7

```

Q8NF17
ID Q8NF17 PRELIMINARY; PRT; 509 AA.
AC Q8NF17;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FJ00385 protein (fragment).
GN FJ00385.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT "the nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK090464; BAC03445.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGC1; 3.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
FT NON TER 1
SQ SEQUENCE 509 AA; 56111 MW; 089498D8076E863C CRC64;

Query Match 38.8%; Score 1126.5; DB 4; Length 509;
Best Local Similarity 63.1%; Pred. No. 7.8e-85;
Matches 229; Conservative 20; Mismatches 57; Indels 57; Gaps 8

Qy 194 LLOENSLY-----TIPKFGSGHLLPFAFLHGNPMLCNCELLYFRWLQDNAENVYVK 247
Db 119 VLOSSGLYSLSVTVFPSSLSGT-----QTYTCN-----VNHKPSNTKVDR 159

Qy 248 QGVYDKAMTSNVASV-----QCDNSDKFPVYKPGKCGPTLGDEGDTLDLYYPEED 299
Db 160 R-VELATPLGDTTHTCPRCEPKSCDTPPCRCPEP-KSCDT-----PPPC 204

Qy 300 TEGDKVARTRVKFPPTKARHTPCPCAPALGAPSVFLFPKPKDTLMISRTPEVTCV 359
Db 205 PRCEPKSCDT-----PPCPRCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCV 255

Qy 360 VVDVSHEDPEVKFNWYDGVGVHNAKTKPREQVNSTYRVVSVLTVLHQDLNKEYKCK 419
Db 256 VVDVSHEDPEVQFKWYDGVGVHNAKTKPREEQNSTFRVVSVLTVLHQDLNKEYKCK 315

Qy 420 VSNKALPVPKTIKATGQGPREFQVYTLPPSREMTKNQVSLTCLVKGFPSDIAVEWS 479
Db 316 VSNKALPAPIKTIKATGQGPREFQVYTLPPSREMTKNQVSLTCLVKGFPSDIAVEW 375

Qy 480 SNGQPENNYTKTPVLDSGSGFFLYSKLTVDKSRWQGNVFPSCSVHMEALHNHYTKSL 539
Db 376 SSGQPENNYNTTPMLDSGSGFFLYSKLTVDKSRWQGNVFPSCSVHMEALHNHYTKSL 435

Qy 540 LSP 542
Db 436 LSP 438

RESULT 8
O35930
ID O35930 PRELIMINARY; PRT; 734 AA.
AC O35930;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Platelet glycoprotein Ib-alpha.
GN GPIBA.
OS Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ;  
RX MEDLINE=98060358; PubMed=9410473;  
RA Ware J., Russell S., Ruggeri Z.;  
RT "Cloning of the murine platelet glycoprotein Ib(alpha) gene  
highlighting species-specific platelet adhesion.";  
RL Blood Cells Mol. Dis. 23:292-301(1997).  
DR EMBL: U91967; AAC33320.1; -;  
DR MGD; MGI:1333744; GpIbA.  
DR InterPro; IPR001114; Hpr\_Serp\_site.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_Typ.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR Pfam; PF00560; LRR; 6.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR0019; LEURICHRPT.  
DR PRINTS; PR0121; PRICHEXTENS.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_Typ; 2.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
DR PROSITE; PS00589; PTS\_HPR\_SER; 1.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
SQ SEQUENCE 734 AA; 80041 MW; 3A32836ASCD5FB5E CRC64;

Query Match 37.0%; Score 1076.5; DB 11; Length 734;  
Best Local Similarity 49.5%; Pred. No. 1.8e-80;

Matches 245; Conservative 52; Mismatches 143; Indels 55; Gaps 10;

QY 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCNKLALPPDLKDTTILHLSNLLY 60  
DB 1 MALLILLFLPSPHSQHTCSISKVTSLEVNCNKLALPADLPADTGILHGENQLG 60  
QY 61 TFSLATMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNLQSLPLGOTLPALT 120  
DB 61 TFSLATMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNLQSLPLGOTLPALT 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLT 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLT 180  
QY 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFR 240  
DB 181 SGULDGLLEDLTLYLQRLNWLRTIPKGFSGTLLPFFVLHANSWYCDCEILYFR 240  
QY 241 ENVYVWKQGVVDVKAMTSNVAQVQDNDSDKFPVYKPGKGTLPGLDGDYDYP 300  
DB 241 ENVYVWKQGVVDVKAMTSNVAQVQDNDSDKFPVYKPGKGTLPGLDGDYDYP 300  
QY 301 EGDVKVATRTVVKFPTKARPHCTCPAPALGAPSVFLFPKPKDTLMISRTPE 360  
DB 298 ---VPATREVKFSTNTKVHT-----THWSLLAA--PSTSQDSQMSISLP 338  
QY 361 VDVSHPDPEVKFWYVDGVEVHNAKTPREEQVNSYRVSVLTJVLQDWLNGEYK 420  
DB 339 ---THK-PTKKQSTFIH-TQSPGFTLLPETMESNPTF-----YSLKL 375  
QY 421 SNKALVPV--IETISKAKQPR-EPQVYTLPPSREEMTKNQVSLTCLVKGFYPS 477  
DB 376 NTVLIPSTPTLEPTSTQATPEPNIQPMLTSTLTPTEHSTTPVPTTILTP 434  
QY 478 WESNGOPENNYKTP 492  
DB 435 -----PTTALITTP 443

RESULT 9

Q28256  
ID Q28256  
AC Q28256;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Glycoprotein Ib.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20011007; PubMed=10544923;  
RA Kenny D., Moratsek P.A., Fahs S.A., Warltier D.C., Montgomery R.R.;  
RT "Cloning and expression of canine glycoprotein Ibalpha.";  
RL Thromb. Haemost. 82:1327-1333(1999).  
DR EMBL: U9489; AAC14361.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_Typ.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR Pfam; PF00560; LRR; 5.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR0019; LEURICHRPT.  
DR PRINTS; PR0121; PRICHEXTENS.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_Typ; 3.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
SQ SEQUENCE 677 AA; 74132 MW; 3798182C4D7F01C0 CRC64;

Query Match 33.4%; Score 970; DB 6; Length 677;  
Best Local Similarity 54.5%; Pred. No. 1.1e-71;

Matches 210; Conservative 33; Mismatches 110; Indels 32; Gaps 5;

QY 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCNKLALPPDLKDTTILHLSNLLY 60  
DB 1 MHLILLMLLLARLCRPEFICEVSKVTSQVEVNCNKGKALPPCLPGDTAILHLAENPLG 60  
QY 61 TFSLATMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNLQSLPLGOTLPALT 120  
DB 61 APSTALLGPTLTAQLHRSQSLTQLQVGMPLRLETLDVSHNRKXSLPSLGRALPALT 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLT 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLT 180  
QY 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFR 240  
DB 181 PGFLEGLGELDTLYLQRLNWLRTIPKGFSGTLLPFFVLHANSWYCDCEILYFR 240  
QY 241 ENVYVWKQGVVDVKAMTSNVAQVQDNDSDKFPVYKPGKGTLPGLDGDYDYP 300  
DB 241 NNVLWKGEVKAATTPNVDSVRCVNWKNVPVHTYQKDCPSPMDGGDMD-YDNYD 299  
QY 301 E-GDKVATRTVVKFPTKARPHCTCP-----PCPA-----PEA 331  
DB 300 KLPGVEAPATHAVSFSTHTKAHTTHMGLLYPTFAYPDHOMAYLSSTLELTEKOT 359  
QY 332 LG--APSVFLPPPKPKDTLMISRT 354  
DB 360 LGPIMPTTPTTPTTLEPTTTP 384

RESULT 10

Q28256  
ID Q28256  
AC Q28256;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Immunoglobulin gamma 1 heavy chain constant region (Fragment).  
 GN IGHCI.

OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;

[1]  
 RP SEQUENCE FROM N.A.

RA Wagner B.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=98383416; PubMed=9717671;  
 RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,  
 RA Leibold W., Radbruch A.

RT "Organization of the equine immunoglobulin heavy chain constant region  
 genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";  
 RL Immunobiology 199,105-119(1998).

DR EMBL; AJ300675; CAC44624.1; -.  
 DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.

DR Pfam; PF00047; Ig\_2.  
 DR SMART; SM00407; IGc1; 2.

DR PROSITE; PS0835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.

FT NON TER 1  
 SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 30.7%; Score 891; DB 6; Length 337;  
 Best Local Similarity 70.1%; Pred. No. 1.5e-65;  
 Matches 157; Conservative 34; Mismatches 31; Indels 2; Gaps 1;

QY 323 CPGCPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 382  
 Db 114 CPKCPAPEALLGGPSVFIAPPNPKDTLMITRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 173

QY 393 NAKTPREQNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALVPVTEKTSKAKGPRE 442  
 Db 174 TATTPKEQNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALVPVTEKTSKAKGPRE 233

QY 443 PQVYLPSPREEMTKNVSCLVLKVGFPSPDIWESNGQP--ENNKTPPVLDSDGS 500  
 Db 234 PQVYLAHPHDELSSKSVTLCKYFPPEINTEWQNGQPELETXYTQAOQSDGS 293

QY 501 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 544  
 Db 294 YFLYSKLSVDRNRWQGGTFTTCGMHEALHNHYTKKNVKNPGK 337

RESULT 11  
 Q8R3V9

ID Q8R3V9 PRELIMINARY; PRT; 469 AA.

AC Q8R3V9;

DT 01-JUN-2002 (TRENBLrel. 21, Created)

DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Hypothetical 52.0 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

[1]  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC024405; AAH24405.1; -.  
 DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_LV.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00406; IGV; 1.

DR PROSITE; PS0835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 28.8%; Score 838; DB 11; Length 469;  
 Best Local Similarity 46.9%; Pred. No. 5.9e-61;  
 Matches 166; Conservative 53; Mismatches 83; Indels 52; Gaps 7;

QY 136 OENSLYT---IPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENVYVMKQGV 252  
 Db 163 QTNQSVTLGCLVKGYFPEPV-----VTWNSGSLSSGVHTFPAVLQ---SDLYTLSSSVTV 215

QY 253 KAMTSNVAQVQDMSKDFPVYKFGKGCPTLGDGDTLDYDYPEEDTEGDKVRAITRV 312  
 Db 216 PSSTWPSQVTCNVA-----HPASSTKVDKTIKTPRDC- 247

QY 313 KEPTKARPHITCPP--CPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEV 370  
 Db 248 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTPKVTCTVVVDISKDDPEV 295

QY 371 KFNWYVDGVEVHNAKTPREQYNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALVPVIE 430  
 Db 296 QFSNFDVDDVEVHTAQTKEPREEQPNSTFRSVSELPIMHODMLNGKEYCKRVNSAFAFPAPIE 355

QY 431 KTIISKAKQPREPOVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 490  
 Db 356 KTIISKTKGRKAPQVYTIIPPKQMAKDKVSLTCTMITDFFPEDITVEMWNGQPAENYKN 415

QY 491 TTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 544  
 Db 416 TQPINDTGGSFYVSKLVNQSNAEAGNTFTCSVLHGLNHNHTEKLSLSPGK 469

RESULT 12  
 Q99LC4

ID Q99LC4 PRELIMINARY; PRT; 463 AA.

AC Q99LC4;

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Similar to RIKEN CDNA 1810060009 gene.  
 GN IGH-4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

[1]  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC003435; AAH03435.1; -.  
 DR HSSP; P01842; 7FAB.

DR MGD; MGI:96446; Igh-4.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_LV.

DR Pfam; PF00047; Ig; 4.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS0835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.

SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 28.7%; Score 835; DB 11; Length 463;  
 Best Local Similarity 46.6%; Pred. No. 1e-60;  
 Matches 165; Conservative 54; Mismatches 83; Indels 52; Gaps 7;

QY 136 OENSLYT---IPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENVYVMKQGV 252  
 Db 157 QTNQSVTLGCLVKGYFPEPV-----VTWNSGSLSSGVHTFPAVLQ---SDLYTLSSSVTV 209

QY 253 KAMTSNVAQVQDMSKDFPVYKFGKGCPTLGDGDTLDYDYPEEDTEGDKVRAITRV 312  
 Db 216 PSSTWPSQVTCNVA-----HPASSTKVDKTIKTPRDC- 247

```

Db 210 PSSWPSSETVTCNVA-----HPASSTKVKKIIPVDC- 241
Qy 313 KPPTKARPHTCPP--CPAPEALGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDDEV 370
Db 242 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTPKVTCCVVDISKDDPEV 289
Qy 371 KFNWYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNKGEYKCKVSNKALPVPIE 430
Db 290 QFSWFVDVVEVHTAQTPREEQENSTRFSVSELPIMHQDWLNKGEYKCKVSNKALPVPIE 349
Qy 431 KTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGYFSPDIAVEWESNGQPPENNYKT 490
Db 350 KTISKTKGRPKAPQVYTIPTPKKQAKDVSLTCLMTDFEFDITVEMQWNGQPAENYKN 409
Qy 491 TPVLDSDGSFFLYSKLTVDKSRWQGNVPSVMEALHNNHYTKSLSPGK 544
Db 410 TQPMIDTDSGYFYKLVNOKSNWEAGNTFTCSVLHGLHNNHTEKLSHSPGK 463

RESULT 13
Q9RI4 PRELIMINARY; PRT; 437 AA.
AC Q9RI4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gamma1 heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; IGH-4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF000047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 28.6%; Score 831; DB 11; Length 437;
Best Local Similarity 46.3%; Pred. No. 2e-60;
Matches 164; Conservative 55; Mismatches 83; Indels 52; Gaps 7;

Qy 196 QENSLYT---IPKGFSGSHLLPEAFLGNPWLNCNCILYFRMLQNAENVYVMKQGVVD 252
Db 131 QTNSTMVTLGCLVKGYPEPTV---VTWNSGSLSSGVHTFPVLQ---SDLYTLSSVTV 183
Qy 253 KAMTSNVASVOCNSDKFPYKYPGKCTPLGDEGDTLDLYPYPEEDTSGDKVRAITRV 312
Db 184 PSSWPSSETVTCNVA-----HPASSTKVKKIIPVDC- 215
Qy 313 KPPTKARPHTCPP--CPAPEALGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDDEV 370
Db 216 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTPKVTCCVVDISKDDPEV 263
Qy 371 KFNWYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNKGEYKCKVSNKALPVPIE 430
Db 264 QFSWFVDVVEVHTAQTPREEQENSTRFSVSELPIMHQDWLNKGEYKCKVSNKALPVPIE 323

```

```

Qy 431 KTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGYFSPDIAVEWESNGQPPENNYKT 490
Db 324 KTISKTKGRPKAPQVYTIPTPKKQAKDVSLTCLMTDFEFDITVEMQWNGQPAENYKN 383
Qy 491 TPVLDSDGSFFLYSKLTVDKSRWQGNVPSVMEALHNNHYTKSLSPGK 544
Db 384 TQPMIDTDSGYFYKLVNOKSNWEAGNTFTCSVLHGLHNNHTEKLSHSPGK 437

RESULT 14
Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Pancreas;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Giasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli K., Bares G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; IGH-1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF000047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 27.8%; Score 808; DB 11; Length 473;
Best Local Similarity 58.3%; Pred. No. 1.8e-58;
Matches 148; Conservative 39; Mismatches 59; Indels 8; Gaps 2;

Qy 295 YPEEDTSGDKVRAITRVVFPPTKARP----HTCCPCPAPEALGAPSVFLPPKPKDTLMI 350
Db 224 HPASSTKVDC---KTEPRVPIQNCPPLKCEPCCAAPDLLGGPSVFFPPKIKDVLMI 279
Qy 351 SRPEVTCVVVDVSHEDDEVKFNWYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODW 410
Db 280 SLSPMTCCVVVDVSDSDPDVQISWFFVNNVEVHTAQTPREEQYNSTRVSVLPIHQDW 339
Qy 411 LNCGEYKCKVSNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGY 470
Db 340 MSGKEFKCKVNNALPSPIETKISKPRGPVAPQVTVLPPPAEEMTKGFSLTGITGL 399

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:15:09 ; Search time 321.846 Seconds  
(without alignment)  
268.288 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPELLLLLLPLSPHPHPIC.....MHEALHNHYTKSLSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	544	23	ABB78234
2	2894	99.6	544	23	ABB78235
3	2881	99.1	544	23	ABB78236
4	2825.5	97.2	531	23	ABB78237
5	2818.5	97.0	531	23	ABB78239
6	2812.5	96.8	531	23	ABB78238
7	2489	85.7	562	21	AA149933
8	2459	84.6	568	21	AA149935
9	1693	58.3	626	22	AA12135

10	1599	55.0	610	17	AA889436	Mutated platelet g
11	1599	55.0	610	18	AAW19201	Platelet glycoprot
12	1596	54.9	610	15	AAAS1116	Platelet glycoprot
13	1592	54.8	610	15	AAAS5664	Mutant platelet g1
14	1586	54.6	302	23	ABB78241	Amino acid sequenc
15	1576.5	54.2	301	23	ABB78240	Amino acid sequenc
16	1553	53.4	293	10	AA91368	45 kDa amino termi
17	1539	53.0	290	23	ABB78243	Amino acid sequenc
18	1533	52.8	290	23	ABB78245	Amino acid sequenc
19	1529	52.6	300	23	ABB78242	Amino acid sequenc
20	1526	52.5	300	23	ABB78244	Amino acid sequenc
21	1236	42.5	313	19	AAW53321	P-selectin ligand
22	1236	42.5	313	20	AAW29766	P-selectin ligand
23	1236	42.5	313	23	AAW77946	Human dimeric p-se
24	1233	42.4	542	23	AAW76357	Pc disintegrin con
25	1232.5	42.4	388	22	AAW50248	Human interleukin-
26	1232	42.4	580	23	AAE23302	Human nectin-4-IgG
27	1232	42.4	651	18	AAW26649	Chimeric receptor
28	1232	42.4	692	18	AAW26650	Chimeric receptor
29	1230.5	42.3	453	14	AAW33311	Humanised Mag11 Ve
30	1230.5	42.3	453	21	AAW85199	Heavy chain amino
31	1230.5	42.3	476	22	AAW49243	Chimeric 4H6 anti-
32	1229.5	42.3	451	20	AAW50031	Human E27 anti-IgE
33	1229.5	42.3	451	20	AAW95659	Human nectin-4-IgG
34	1229.5	42.3	451	20	AAW95661	Chimeric receptor
35	1229.5	42.3	451	20	AAW95663	Mus musculus anti-
36	1229.5	42.3	451	21	AAW07473	Mus musculus anti-
37	1229.5	42.3	451	21	AAW85201	Mus musculus anti-
38	1229.5	42.3	451	22	AAW47412	Amino acid sequenc
39	1229.5	42.3	451	22	AAW47088	Light chain amino
40	1229.5	42.3	451	22	AAW76948	E27 anti-IgE antib
41	1229.5	42.3	451	22	AAW76950	Anti-IgE antibody,
42	1229.5	42.3	451	22	AAW76952	Full length heavy
43	1226.5	42.2	502	22	AAW37109	Full length heavy
44	1226	42.2	528	23	AAW79654	Concatermic immun
45	1225	42.2	694	21	AAW92201	Human soluble PSG1
						Fusion polypeptide

#### ALIGNMENTS

#### RESULT 1

ABB78234	ID	ABB78234 standard; Protein; 544 AA.
XX	AC	ABB78234;
XX	DT	25-NOV-2002 (first entry)
XX	DE	Amino acid sequence of GPLB302-Ig fusion protein.
XX	KW	Glycoprotein 1B-alpha; GPLB; immunoglobulin; Ig; platelet adherence;
KW	KW	leukocyte; platelet activation; ischaemic heart disease;
KW	KW	acute myocardial infarction; angina; stroke; venous thrombosis; atherosclerosis;
KW	KW	arterial thrombosis; angina; stroke; venous thrombosis; atherosclerosis;
XX	OS	Thrombosis; angioplasty; restenosis.
XX	XX	Synthetic.
XX	PH	Key
XX	FT	Misc-difference 500
XX	FT	/note= "Ser encoded by CCC"
XX	PN	WO200263003-A2.
XX	XX	15-AUG-2002.
XX	XX	06-FEB-2002; 2002WO-US03549.
XX	XX	06-FEB-2001; 2001US-266838P.
XX	XX	(GEMY ) GENETICS INST LLC.
XX	XX	

PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78661.  
 XX  
 XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis  
 XX  
 XX Claim 5; Page 2; 45pp; English.  
 PS  
 XX The present sequence represents a fusion protein of glycoprotein  
 CC 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 XX Sequence 544 AA;  
 Query Match 100.0%; Score 2906; DB 23; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-216;  
 Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MPELLLLLLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
 Db 1 MPELLLLLLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLVQDGTLPVLGTLDSLHNQSLPLGQTLPALTV 120  
 Db 61 TFSLATLMPYTRLTQNLNDRCELTKLVQDGTLPVLGTLDSLHNQSLPLGQTLPALTV 120  
 QY 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
 Db 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
 QY 181 AGLINGLENLDTLLQENSLLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 Db 181 AGLINGLENLDTLLQENSLLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 QY 241 ENVYVWKQGVVDVKAMTSNVASVQCNSDKFPVYKYPGKCPGLGEGDTLDLYDYPEDT 300  
 Db 241 ENVYVWKQGVVDVKAMTSNVASVQCNSDKFPVYKYPGKCPGLGEGDTLDLYDYPEDT 300  
 QY 301 EGDKVRATRTVVKFPTKARPHCTPCPCPAEALGAPSVLFPKPKDTLMISRTPEVTCV 360  
 Db 301 EGDKVRATRTVVKFPTKARPHCTPCPCPAEALGAPSVLFPKPKDTLMISRTPEVTCV 360  
 QY 361 VDVSHEDEPVKNWYVDGVEVNAKTKPREQYNSTYRVVSVLTVLHQDLNGKEYKCKV 420  
 Db 361 VDVSHEDEPVKNWYVDGVEVNAKTKPREQYNSTYRVVSVLTVLHQDLNGKEYKCKV 420  
 QY 421 SNKALVPVTEKTSKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGYPSDIAVENES 480  
 Db 421 SNKALVPVTEKTSKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGYPSDIAVENES 480  
 QY 481 NGQPENNYKTTTPVLSDSGSFYLSKLTVDKSRWQQGNVFCGVMHEALHNNHYTKS-SL 540  
 Db 481 NGQPENNYKTTTPVLSDSGSFYLSKLTVDKSRWQQGNVFCGVMHEALHNNHYTKS-SL 540  
 QY 541 SPQK 544  
 Db 541 SPQK 544

RESULT 2

ABB78235

ID ABB78235 standard; Protein; 544 AA.

XX

AC ABB78235;

XX 25-NOV-2002 (first entry)  
 DT Amino acid sequence of GPIb302/2A-Ig fusion protein.  
 XX  
 DE Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 XX leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 XX thrombosis; angioplasty; restenosis.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 500 /note= "Ser encoded by CCC"  
 FT  
 FT WO200263003-A2.  
 XX  
 XX 15-AUG-2002.  
 PD  
 XX 06-FEB-2002; 2002WO-US03549.  
 PF  
 XX 06-FEB-2001; 2001US-266838P.  
 PR  
 XX (GEMY ) GENETICS INST LLC.  
 PA  
 XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 PI WPI; 2002-657537/70.  
 XX N-PSDB; ABQ78662.  
 DR  
 XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis  
 XX  
 XX Claim 20; Page 3; 45pp; English.  
 PS  
 XX The present sequence represents a fusion protein of glycoprotein  
 CC 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 CC  
 XX Sequence 544 AA;  
 Query Match 99.6%; Score 2894; DB 23; Length 544;  
 Best Local Similarity 99.6%; Pred. No. 4.1e-215;  
 Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MPELLLLLLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
 Db 1 MPELLLLLLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLVQDGTLPVLGTLDSLHNQSLPLGQTLPALTV 120  
 Db 61 TFSLATLMPYTRLTQNLNDRCELTKLVQDGTLPVLGTLDSLHNQSLPLGQTLPALTV 120  
 QY 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
 Db 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
 QY 181 AGLINGLENLDTLLQENSLLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 Db 181 AGLINGLENLDTLLQENSLLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 QY 241 ENVYVWKQGVVDVKAMTSNVASVQCNSDKFPVYKYPGKCPGLGEGDTLDLYDYPEDT 300  
 Db 241 ENVYVWKQGVVDVKAMTSNVASVQCNSDKFPVYKYPGKCPGLGEGDTLDLYDYPEDT 300

CC	angioplasty-related restenosis.			
XX	Sequence	544	AA;	
QQ				
	Query Match	99.1%	Score 2881;	DB 23; Length 544;
	Best Local Similarity	99.3%;	Pred. No. 4.2e-214;	
	Matches 540; Conservative	1;	Mismatches 3;	Indels 0; Gaps 0;
QY	1	MPLLLLLLLLPSPLPHPHPICEVSKVASHLEVNC	KRNLTALPPDLPKDPTILHLSENLLY	60
DB	1	MPLLLLLLLLPSPLPHPHPICEVSKVASHLEVNC	KRNLTALPPDLPKDPTILHLSENLLY	60
QY	61	TSLSLATIMXYTRLTQNLNRCCELTKLQVDGTL	PVLGTLDLSHNQLQSLPVLGQTLPLATV	120
DB	61	TSLSLATIMPYTRLTQNLNRCCELTKLQVDGTL	PVLGTLDLSHNQLQSLPVLGQTLPLATV	120
QY	121	LDVSFNRLTSLPGLARGLGELOELYKGNELKTL	PPGLTPTPKLEKLSLANNNLTELP	180
DB	121	LDVSFNRLTSLPGLARGLGELOELYKGNELKTL	PPGLTPTPKLEKLSLANNNLTELP	180

[illegible]

PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-US03549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GEM) ) GENETICS INST LLC.  
 XX  
 PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78664.  
 XX  
 PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 3; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 97.2%; Score 2825.5; DB 23; Length 531;  
 Best Local Similarity 97.6%; Pred. No. 7.9e-210;  
 Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
 QY 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTLHLSENLY 60  
 DB 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTLHLSENLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 QY 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180  
 DB 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180  
 QY 181 AGLLNGLENLDTLLQENSLYTIIPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 DB 181 AGLLNGLENLDTLLQENSLYTIIPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240  
 QY 241 ENYVWKQGVYDVVAMTSNVASVOCDSNDKPPVYKPGKCGPTLGDGDDTLDYYPEEDT 300  
 DB 241 ENYVWKQGVYDVVAMTSNVASVOCDSNDKPPVYKPGKCGPTLGDGDDTLDYYPEEDT 300  
 QY 301 EGDVRAVTRIVKFTPKARPHTCPPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
 DB 301 EGDVRAVTRIVKFTPKARPHTCPPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
 QY 361 VDVSHEDPEVKNYVDGVEVHNAKTPREQYNSTRVSVLTVLHQDLNGKEYKCV 420  
 DB 361 VDVSHEDPEVKNYVDGVEVHNAKTPREQYNSTRVSVLTVLHQDLNGKEYKCV 420  
 QY 421 SNKALPVPIDKTSKAGQPREQVYTLPPSREMTKNQVSLTCLVKGFYFSPDIAVWES 480  
 DB 421 SNKALPVPIDKTSKAGQPREQVYTLPPSREMTKNQVSLTCLVKGFYFSPDIAVWES 480  
 QY 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCWMEALHNNHYTKSL 540  
 DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCWMEALHNNHYTKSL 540  
 QY 541 SPKG 544  
 |||||

Db 528 SPKG 531  
 RESULT 5  
 ABB78239  
 ID ABB78239 standard; Protein; 531 AA.  
 XX  
 AC ABB78239;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of GPIIb290/IA-Ig fusion protein.  
 XX  
 KW Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 253 /note= "Val encoded by GCG"  
 FT Misc-difference 487 /note= "Ser encoded by CCC"  
 XX  
 PN WO200263003-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-US03549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GEM) ) GENETICS INST LLC.  
 XX  
 PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78666.  
 XX  
 PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 4; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 97.0%; Score 2818.5; DB 23; Length 531;  
 Best Local Similarity 97.4%; Pred. No. 2.7e-209;  
 Matches 530; Conservative 0; Mismatches 1; Indels 13; Gaps 1;  
 QY 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTLHLSENLY 60  
 DB 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTLHLSENLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 QY 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180

Db 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGGLTTPPKLEKLSLANNLTLP 180  
Qy 181 AGLNGLENLDTLLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Qy 241 ENVYVWKQGVVDKAMTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYVWKQGVVDKAMTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKVRATRTVVKFPTKARPHTCPCPAPEALGAPSVLFPKPKDTLMISRTPEVTCV 360  
Db 301 EGDKV-----RPHTCPCPAPEALGAPSVLFPKPKDTLMISRTPEVTCV 347  
Qy 361 VDVSHEDEPKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYCKV 420  
Db 348 VDVSHEDEPKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYCKV 407  
Qy 421 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAWEMES 480  
Db 408 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAWEMES 467  
Qy 481 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKLSL 540  
Db 468 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKLSL 527  
Qy 541 SPKG 544  
Db 528 SPKG 531

RESULT 6  
ABB78238  
ID ABB78238 standard; Protein; 531 AA.

AC ABB78238;  
DT 25-NOV-2002 (first entry)  
XX Amino acid sequence of GPlb290/2V-Ig fusion protein.

XX Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.

XX Synthetic.  
XX Key Location/Qualifiers  
FT Misc-difference 487 /note= "Ser encoded by CCC"

XX WO200263003-A2.  
XX 15-AUG-2002.  
XX 06-FEB-2002; 2002WO-US03549.  
XX 06-FEB-2001; 2001US-266838P.  
XX (GEMY ) GENETICS INST LLC.  
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
XX N-PSDB; ABQ78665.

XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
XX disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX

PS Claim 20; Page 3-4; 45pp; English.  
XX The present sequence represents a fusion protein of glycoprotein  
CC 1B-alpha (GPlb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX  
SQ Sequence 531 AA;

Query Match 96.8%; Score 2812.5; DB 23; Length 531;  
Best Local Similarity 97.2%; Pred. No. 8e-209; 1; Indels 13; Gaps 1;  
Matches 529; Conservative 1; Mismatches 1;

Qy 1 MPLLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDPKDTTILHLSENLLY 60  
Db 1 MPLLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDPKDTTILHLSENLLY 60  
Qy 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLGOTLPALT 120  
Db 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLGOTLPALT 120  
Qy 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGGLTTPPKLEKLSLANNLTLP 180  
Db 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGGLTTPPKLEKLSLANNLTLP 180  
Qy 181 AGLNGLENLDTLLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Qy 241 ENVYVWKQGVVDKAMTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYVWKQGVVDKAMTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKVRATRTVVKFPTKARPHTCPCPAPEALGAPSVLFPKPKDTLMISRTPEVTCV 360  
Db 301 EGDKV-----RPHTCPCPAPEALGAPSVLFPKPKDTLMISRTPEVTCV 347  
Qy 361 VDVSHEDEPKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYCKV 420  
Db 348 VDVSHEDEPKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYCKV 407  
Qy 421 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAWEMES 480  
Db 408 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAWEMES 467  
Qy 481 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKLSL 540  
Db 468 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKLSL 527  
Qy 541 SPKG 544  
Db 528 SPKG 531

RESULT 7  
AAY49933  
ID AAY49933 standard; Protein; 562 AA.

XX AAY49933;  
XX 01-FEB-2000 (first entry)  
XX Human glycoprotein 1b/mouse IgG1Fc chimeric protein.  
XX Glycoprotein 1b; glycoallidin; detection; antithrombotic; binding;  
KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;  
KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.  
XX

OS Chimeric - Homo sapiens.  
OS Chimeric - Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..16  
FT /label= signal  
FT Protein 17..562  
FT /label= human glycoprotein Ib/mouse\_IgG1Fc\_chimeric  
FT protein  
XX  
XX W09954360-A1.  
XX  
XX 28-OCT-1999.  
XX 13-JAN-1999; 99WO-JP00089.  
XX 23-APR-1998; 98JP-0113962.  
XX (AJIN ) AJINOMOTO CO INC.  
XX Fukuchi N, Furaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;  
XX Tanaka A, Shinozaki J, Jojima Y;  
XX MPI; 2000-013233/01.  
XX N-PSDB; AA235701.  
XX  
XX Direct quantitative detection of glycoallidin with immobilized von  
XX Willebrand factor to bond with chimeric protein via inhibiting  
XX Glycoprotein Ib binding, for diagnosis of thrombotic diseases and  
XX screening anti-thrombotic substances -  
XX  
XX Example 1; Page 70-71; 83pp; Japanese.  
XX  
XX A method has been developed for the detection of binding between the von  
XX Willebrand factor and glycoprotein Ib or of the binding inhibition, in  
XX which the von Willebrand factor immobilised in a reactor reacts with  
XX glycoprotein Ib in the presence of a binding inducer to promote binding  
XX between the von Willebrand factor and glycoprotein Ib. This method is  
XX for the detection of glycoallidin as a means of thrombotic disease  
XX diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
XX for screening substances with anti-thrombotic activity for the  
XX prevention and treatment of thrombotic diseases. The method is direct,  
XX convenient and quantitative, with reproducibility, and there is no need  
XX to construct a monoclonal antibody for the assay. The present sequence  
XX is a human glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric  
XX protein from the present invention.  
XX  
XX Sequence 562 AA;  
SQ  
Query March 85.7%; Score 2489; DB 21; Length 562;  
Best Local Similarity 81.7%; Pred. No. 8.8e-184;  
Matches 463; Conservative 39; Mismatches 37; Indels 28; Gaps 4;  
QY 1 MPILLLLLLLPSLPHHPICEVSKVASHLEWNCCKNLTPALPPDLPKDTTLHLSENLLY 60  
DB 1 MPILLLLLLLPSLPHHPICEVSKVASHLEWNCCKNLTPALPPDLPKDTTLHLSENLLY 60  
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSLHNLQSLPLGQTLPAJTV 120  
DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSLHNLQSLPLGQTLPAJTV 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLITELP 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLITELP 180  
QY 181 AGLLNGLENLDTLLQNSLYTTPKPGFSGHLLPFAFLHGNPNWLCNCEILYFRWLQDNA 240  
DB 181 AGLLNGLENLDTLLQNSLYTTPKPGFSGHLLPFAFLHGNPNWLCNCEILYFRWLQDNA 240  
QY 241 ENVYVWKQGVVDVKAWTSNVASVOCNDSKFPVVKYKPGKCGPTLGDGSDTLDLYDYPEEDT 300  
DB 241 ENVYVWKQGVVDVKAWTSNVASVOCNDSKFPVVKYKPGKCGPTLGDGSDTLDLYDYPEEDT 300

QY 301 EGDVRAIRTRTVVKFPTKARPHT-----CPP--CPAPEALGAPSV 337  
DB 301 EGDVRAIRTRTVVKFPTKA--HTTPWGLFYSWSPASLDVPRDCGCKPCICTVPE--VSSV 355  
QY 338 FLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFMYVDGVENVHNAKTKPREEQNSTY 397  
DB 356 FIFPKPKDVLITITLTPKVTCCVVVDISKDDPEVQFSWFVDVDEVHTAQTQPREQFNSTF 415  
QY 398 RVYSVLTVLHODWLNKGEYKCKVSNKALPVPIEKTISKAKGQPREPQVYTLPPSREEMTK 457  
DB 416 RVSSELPIMHQDWLNGKEFKCRVNSAFAPIEKTISKTKGRKPAQVTTIPPKQMAK 475  
QY 458 NOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQG 517  
DB 476 DKVSLTCTMIDTFPPEDITVEMQWNGQPAENYKNTQPIMTDGSYFYVYSKLVQKSNWEAG 535  
QY 518 NVFSCSVMEALHNRHTQKSLSPGK 544  
DB 536 NTFCTSVLHEGLNHNHTKSLSHSPGK 562  
RESULT 8  
AAI49935  
ID AAY49935 standard; Protein; 568 AA.  
XX  
XX AAY49935;  
XX  
XX 01-FEB-2000 (first entry)  
XX Human glycoprotein Ib/mouse IgG2aFc chimeric protein #2.  
XX  
XX Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;  
XX von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;  
XX thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.  
XX Chimeric - Homo sapiens.  
XX Chimeric - Mus musculus.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..16  
XX /label= signal  
XX Protein 17..568  
XX /label= Human glycoprotein Ib/mouse\_IgG2aFc\_chimeric  
XX protein  
XX W09954360-A1.  
XX  
XX 28-OCT-1999.  
XX  
XX 13-JAN-1999; 99WO-JP00089.  
XX  
XX 23-APR-1998; 98JP-0113962.  
XX (AJIN ) AJINOMOTO CO INC.  
XX Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;  
XX Tanaka A, Shinozaki J, Jojima Y;  
XX MPI; 2000-013233/01.  
XX N-PSDB; AA235706.  
XX  
XX Direct quantitative detection of glycoallidin with immobilized von  
XX Willebrand factor to bond with chimeric protein via inhibiting  
XX glycoprotein Ib binding, for diagnosis of thrombotic diseases and  
XX screening anti-thrombotic substances -  
XX  
XX Example 1; Page 77-79; 83pp; Japanese.  
XX  
XX A method has been developed for the detection of binding between the von  
XX Willebrand factor and glycoprotein Ib or of the binding inhibition, in  
XX which the von Willebrand factor immobilised in a reactor reacts with  
XX glycoprotein Ib in the presence of a binding inducer to promote binding  
XX between the von Willebrand factor and glycoprotein Ib. This method is  
XX for the detection of glycoallidin as a means of thrombotic disease  
XX diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
XX for screening substances with anti-thrombotic activity for the  
XX prevention and treatment of thrombotic diseases. The method is direct,  
XX convenient and quantitative, with reproducibility, and there is no need  
XX to construct a monoclonal antibody for the assay. The present sequence  
XX is a human glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric  
XX protein from the present invention.  
XX  
XX Sequence 562 AA;  
SQ

CC for the detection of glykocallidin as a means of thrombotic disease  
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
CC for screening substances with anti-thrombotic activity for the  
CC prevention and treatment of thrombotic diseases. The method is direct,  
CC convenient and quantitative, with reproducibility, and there is no need  
CC to construct a monoclonal antibody for the assay. The present sequence  
CC represents a human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc  
CC chimeric protein from the present invention.  
XX  
XX  
SQ Sequence 568 AA;  
  
Query Match 84.6%; Score 2459; DB 21; Length 568;  
Best Local Similarity 81.6%; Pred. No. 1.9e-181;  
Matches 465; Conservative 28; Mismatches 49; Indels 28; Gaps 3;  
  
Qy 1 MPULLLLLLLPSPLPHPICEVSKVASHLEVNCCKNLTPALPDLPKDTTILHLENLY 60  
Db 1 MPULLLLLLLPSPLPHPICEVSKVASHLEVNCCKNLTPALPDLPKDTTILHLENLY 60  
  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLGQTLPALTV 120  
  
Qy 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
  
Qy 181 AGLNGLENLDTLLQENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLQENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
  
Qy 241 ENVVVMKQGVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPDGLDGDGTDLDYYPEDT 300  
Db 241 ENVVVMKQGVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPDGLDGDGTDLDYYPEDT 300  
  
Qy 301 EGDVRAIRTVVKKFPTKARPH-----CPP--CPAPEALGA 334  
Db 301 EGDVRAIRTVVKKFPTKA--HTPWGLFYSWSTASLDEPRGPTIKPCPPCKPAPNLGG 358  
  
Qy 335 PSVLPFPKPDLMISHTPEVTCVVDVSHEDVEVFNWTVGVVHNAKTKPREQYN 394  
Db 359 PSVFIFPKIKDVLMSISPIVTCVVDVSDDDPDQVTSFVNNVHTAQTQTHREDYN 418  
  
Qy 395 STYVWSVLTVLHODLNGKVKCNKALPVPIEKTISKAGQREPQVYTLPPSREE 454  
Db 419 STLKWSALPIQHQDWSGKFKCKVKNKDLPAPIERTISKPGSVRAQVYVLPPEEE 478  
  
Qy 455 MTKNQSLTCLVKGFYPSDIAVWESNGQPNYKTTTPPVLDSDGSPFLYSLKLVDSRW 514  
Db 479 MTKKQVTLTCVTFMPEDIYVWVWNNKGTLYNTPEVLDSDGSPFMYSKLVRKQW 538  
  
Qy 515 QQGVFSCSVMEALHNNHYTKSLSLSPK 544  
Db 539 VERNYSQSVVHEGLHNNHTKSFSTRTEK 568

RESULT 9  
AAE12135  
ID AAE12135 standard; Protein; 626 AA.  
XX  
AC AAE12135;  
XX  
XX  
DT 03-JAN-2002 (first entry)  
XX Human glycoprotein Ib (platelet) alpha (GPIbA).  
DE  
DE Human; haplotyping; glycoprotein Ib (platelet) alpha protein; GPIbA;  
KW Bernard-Soulier syndrome; platelet-type von Willebrand disease; HIV;  
KW Alzheimer's disease; human immunodeficiency virus; SNP;  
KW single nucleotide polymorphism; chromosome 17pter-p12.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FH Misc-difference 72  
FT /note= "Arg at this position is replaced with His  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 86  
FT /note= "Leu at this position is replaced with Phe  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 161  
FT /note= "Thr at this position is replaced with Met  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 592  
FT /note= "Arg at this position is replaced with His  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 624  
FT /note= "His at this position is replaced with Arg  
FT due to single nucleotide polymorphism (SNP)"  
FT  
XX  
XX  
PN W0200175065-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US10671.  
XX  
XX 03-APR-2000; 2000US-194341P.  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
XX Bentivegna SC, Choi JY, Kliem SE, Koshy B, Parks KE;  
XX  
XX WPI; 2001-626427/72.  
DR N-PSDB; AAD20681.  
DR  
XX  
XX New haplotypes of the glycoprotein Ib platelet alpha polypeptide gene  
PT are useful for diagnosis and drug discovery for treating Bernard  
PT Soulier syndrome, platelet-type von Willebrand disease, HIV and  
PT Alzheimer's disease -  
XX  
XX Claim 29; Fig 3; 66pp; English.  
XX  
XX The invention relates to methods for haplotyping glycoprotein Ib  
CC (platelet) alpha polypeptide (GPIbA) gene of an individual. The  
CC method involves determining if the individual has one of the GPIbA  
CC haplotypes or haplotype pairs. The methods of the invention are  
CC useful for disease diagnosis and in the discovery and development  
CC of drugs for treating diseases associated with GPIbA activity e.g.  
CC Bernard-Soulier syndrome, platelet-type von Willebrand disease, HIV  
CC and Alzheimer's disease. The present sequence is human GPIbA protein.  
CC GPIbA gene is located on chromosome 17pter-p12.  
XX  
SQ Sequence 626 AA;  
  
Query Match 58.3%; Score 1693; DB 22; Length 626;  
Best Local Similarity 87.4%; Pred. No. 3.6e-122;  
Matches 333; Conservative 4; Mismatches 18; Indels 26; Gaps 5;  
  
Qy 1 MPULLLLLLLPSPLPHPICEVSKVASHLEVNCCKNLTPALPDLPKDTTILHLENLY 60  
Db 1 MPULLLLLLLPSPLPHPICEVSKVASHLEVNCCKNLTPALPDLPKDTTILHLENLY 60  
  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLGQTLPALTV 120  
  
Qy 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
  
Qy 181 AGLNGLENLDTLLQENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLQENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
  
Qy 241 ENVVVMKQGVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPDGLDGDGTDLDYYPEDT 300  
Db 241 ENVVVMKQGVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPDGLDGDGTDLDYYPEDT 300



Query Match	55.0%	Score 1599;	DB 17;	Length 610;
Best Local Similarity	86.3%	Pred. No. 6.4e-115;		
Matches 315;	Conservative 4;	Mismatches 20;	Indels 26;	Gaps 5;

QY	17	HPICEVSKVASHLEVNCDEKENTALPPDLPKQTTHLHSENLLYTFSLATLMPYTRLTQL	76
DB	1	HPICEVSKVASHLEVNCDEKENTALPPDLPKQTTHLHSENLLYTFSLATLMPYTRLTQL	60
QY	77	NLDRCELTKLVQDGTLPVLGTLDSLHNQLSPLLGQTLPAITVDVSFNRLTSLPLGAL	136
DB	61	NLDRCELTKLVQDGTLPVLGTLDSLHNQLSPLLGQTLPAITVDVSFNRLTSLPLGAL	120
QY	137	RGLGELOELYLKNEUKETLPPGLTTPPKLEKLANNLTTELPAGLINGLENLDTLLQ	196
DB	121	RGLGELOELYLYKNELKTLPGLTTPPKLEKLANNLTTELPVGLNGLENLDTLLQ	180
QY	197	ENSLYTIPIKPGFGSHLLPFAFLHGNPWLCNCEILYFRWLQDNAENVYWKQGVDRKMT	256
DB	181	ENSLYTIPIKPGFGSHLLPFAFLHGNPWLCNCEILYFRWLQDNAENVYWKQGVDRKMT	240
QY	257	SNVASVQCDSNDKEFPVKYPKGCGPTLGDGGTDLDLYYPEEDTEGDKVRATRTVVKKPT	316
DB	241	SNVASVQCDSNDKEFPVKYPKGCGPTLGDGGTDLDLYYPEEDTEGDKVRATRTVWKPFT	300
QY	317	KARPHICP-----PCPA---PEALGAPSVFLFPPK--PKDTL-----MISR	352
DB	301	KA--HTPWGLFYSMWTASIDSQMPSSLHPTQSEKQTTFFPRWTNFLLHMESITFSK	358
QY	353	TPEVT 357	
DB	359	TPKST 363	

RESULT 11  
AAWI8201  
ID AAWI8201 standard; protein; 610 AA.  
XX AC AAWI8201;  
XX DT 25-MAR-2003 (updated)  
XX DT 18-AUG-1997 (first entry)  
XX DE Platelet glycoprotein Ib alpha naturally-occurring wild-type.  
XX KW GpIb alpha; recombinant polypeptide; antithrombotic;  
XX KW platelet adhesion; platelet aggregation; thrombus formation;  
XX KW platelet-type von Willebrand disease.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
FT Region 36..200 /label= Leucine\_rich  
FT Region 220..310 /label= Hinge  
FT FT /note= "A major binding site for von Willebrand factor"  
FT Region 228..238 /note= "Preferred sites for mutations that result  
FT FT in a mutant polypeptide having more reactivity  
FT FT with von Willebrand factor"

FT FT Misc-difference 233 /note= "Preferred mutation; substitution of Val for Gly"  
FT FT Region 310..420 /label= Serine/threonine-rich  
FT XX US5624817-A.  
PN XX  
XX PD 29-APR-1997.  
XX PF 28-APR-1994; 94US-0234265.  
XX PR 07-OCT-1991; 91US-0770969.



Db 301 KA--HTTPWCLFYSWSTASLDQMPSSLHPTQESTKEQTTPPRWTPNFTLHVESITFSK 358  
 QY 353 TPEVT 357  
 Db 359 TPKST 363

RESULT 13  
 AAR5664  
 ID AAR5664 standard; protein; 610 AA.  
 XX  
 AC AAR5664;  
 DT 25-MAR-2003 (updated)  
 DT 23-SEP-1994 (first entry)  
 XX  
 DE Mutant platelet glycoprotein Ib alpha.  
 XX  
 KW platelet; glycoprotein; imaging; thrombolytic agent;  
 KW tissue plasminogen activator; tPA; pro-urokinase; urokinase;  
 KW streptokinase; Bernard-Soulier disease; thrombus; aggregation;  
 KW anisoylated plasminogen-streptokinase activator complex; adhesion;  
 KW inhibition.  
 XX  
 OS Homo sapiens.  
 XX  
 PN U55298239-A.  
 XX  
 PD 29-MAR-1994.  
 XX  
 PF 15-JAN-1992; 92US-0821717.  
 XX  
 PR 07-OCT-1991; 91US-0770968.  
 PR 15-JAN-1992; 92US-0821717.  
 XX  
 PA (UYN ) UNIV NEW YORK STATE RES FOUND.  
 XX  
 PI Cunningham D, Finch CN, Lyle VA, Miller JL;  
 XX  
 DR WPI; 1994-100287/12.  
 XX  
 PT Platelet glycoprotein Ib alpha with an amino acid subunit at  
 P position 57 - has reduced reactivity with Von Willebrand factor,  
 P and can be used to inhibit platelet aggregation and inhibition  
 XX  
 PS Claim 1; Page ?; 20pp; English.  
 XX  
 CC A substitution in platelet glycoprotein Ib alpha (Leucine 57 to  
 CC Phenylalanine) underlies a form of Bernard-Soulier disease. The  
 CC mutated glycoprotein can be used in compositions to inhibit  
 CC platelet aggregation/adhesion. The glycoprotein may be labelled and  
 CC used as an imaging agent and may also be bound to a thrombolytic  
 CC agent, preferably tissue plasminogen activator (tPA),  
 CC (pro) urokinase, streptokinase, anisoylated plasminogen-streptokinase  
 CC activator complex, tPA analogues or a protease, allowing localisation  
 CC of the thrombolytic agent to a thrombus.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX

SQ Sequence 610 AA;  
 Query Match 54.8%; Score 1592; DB 15; Length 610;  
 Best Local Similarity 85.8%; Pred. No. 2.2e-114;  
 Matches 313; Conservative 5; Mismatches 21; Indels 26; Gaps 5;

Qy 17 HPICEVSKVASHLEVNCCKNLTALPPDLPKDTTILHSENLLYTSLATMPYTRTQL 76  
 Db 1 HPICEVSKVASHLEVNCCKNLTALPPDLPKDTTILHSENLLYTSLATMPYTRTQL 60

Qy 77 NLDRCETLKLQVDTGLTVLGTDLDSHNLQSLPLGLGTLFALTVDSFNRLTSLPLGL 136  
 Db 61 NLDRCETLKLQVDTGLTVLGTDLDSHNLQSLPLGLGTLFALTVDSFNRLTSLPLGL 120

QY 137 RGLGELQELKGNELKXLTLPGLLTPTPKLEKLSLANNLTLPAGLNGLENDTLILQ 196  
 Db 121 RGLGELQELKGNELKXLTLPGLLTPTPKLEKLSLANNLTLPVGLNGLENDTLILQ 180  
 QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENYVWKQGVVKAMT 256  
 Db 181 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENYVWKQGVVKAMT 240  
 QY 257 SNVASVQCDNSDKFPYKYKGCPTLGDGDDTLVDYYPEEDTEGDKVRAATVVKFPT 316  
 Db 241 SNVASVQCDNSDKFPYKYKGCPTLGDGDDTLVDYYPEEDTEGDKVRAATVVKFPT 300  
 QY 317 KARPHTCP-----PCPA---PEALGAPSVFLFPK--PKOTL-----MIS 352  
 Db 301 KA--HTTPWGLFYSWSTASLDQMPSSLHPTQESTKEQTTPPRWTPNFTLHVESITFSK 358  
 QY 353 TPEVT 357  
 Db 359 TPKST 363

RESULT 14  
 ABB78241  
 ID ABB78241 standard; Protein; 302 AA.  
 XX  
 AC ABB78241;  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of glycoprotein IB-alpha polypeptide GPIb302/2A.  
 XX  
 KW Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200263003-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-US03549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GENY ) GENETICS INST LLC.  
 XX  
 PI Shaw GN, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 XX  
 PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Disclosure; Page 10; 45pp; English.  
 XX  
 CC ABB78240-45 represent glycoprotein IB-alpha (GPIb) polypeptides, which  
 CC are used to produce fusion proteins with an immunoglobulin (Ig)  
 CC polypeptide. The fusion proteins inhibit the adherence of platelets to  
 CC leukocytes. The fusion polypeptides are useful for treating a disorder  
 CC associated with platelet activation e.g. ischaemic heart disease, acute  
 CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,  
 CC arterial thrombosis or unstable angina. They can also be used to treat  
 CC vascular conditions associated with vascular inflammation, thrombosis,  
 CC and angioplasty-related restenosis.  
 XX

SQ Sequence 302 AA;  
 Query Match 54.6%; Score 1586; DB 23; Length 302;  
 Best Local Similarity 99.3%; Pred. No. 2.5e-114;

Matches 300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 HPICEVSVASHLEVNCNDRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSVASHLEVNCNDRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDGTLPLVLGTLDSLHNQQLSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 136  
DB 61 NLDRCCLTKLQVDGTLPLVLGTLDSLHNQQLSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRWLQDQNAENVYVWKQGVVDVKAMT 256  
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRWLQDQNAENVYVWKQGVVDVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPGKCPGLGDEGDTLDLYYYPBEDTEGDKVRATRTVVVFPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPGKCPGLGDEGDTLDLYYYPBEDTEGDKVAATATVVKFPT 300  
QY 317 KA 318  
DB 301 KA 302

RESULT 15  
ABB78240  
ID ABB78240 standard; Protein; 301 AA.  
AC ABB78240;  
DT 25-NOV-2002 (first entry)  
XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb302.  
DE Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX  
OS Unidentified.  
XX  
XX WO200263003-A2.  
XX  
XX 15-AUG-2002.  
XX  
XX 06-FEB-2002; 2002WO-US03549.  
XX  
XX 06-FEB-2001; 2001US-266838P.  
XX  
XX (GENY ) GENETICS INST LLC.  
XX  
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
XX  
XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Disclosure; Page 10; 45pp; English.  
XX  
XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which  
CC are used to produce fusion proteins with an immunoglobulin (Ig)  
CC polypeptide. The fusion proteins inhibit the adherence of platelets to  
CC leukocytes. The fusion polypeptides are useful for treating a disorder  
CC associated with platelet activation e.g. ischaemic heart disease, acute  
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,  
CC arterial thrombosis or unstable angina. They can also be used to treat  
CC vascular conditions associated with vascular inflammation, thrombosis,

CC and angioplasty-related restenosis.

XX  
SQ Sequence 301 AA;  
Query Match 54.2%; Score 1576.5; DB 23; Length 301;  
Best Local Similarity 99.3%; Pred. No. 1.3e-113;  
Matches 300; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 17 HPICEVSVASHLEVNCNDRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSVASHLEVNCNDRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDGTLPLVLGTLDSLHNQQLSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 136  
DB 61 NLDRCCLTKLQVDGTLPLVLGTLDSLHNQQLSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRWLQDQNAENVYVWKQGVVDVKAMT 256  
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRWLQDQNAENVYVWKQGVVDVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPGKCPGLGDEGDTLDLYYYPBEDTEGDKVRATRTVVVFPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPGKCPGLGDEGDTLDLYYYPBEDTEGDKVRATRT -WKFPPT 299  
QY 317 KA 318  
DB 300 KA 301

Search completed: October 14, 2003, 06:30:23  
Job time : 327.846 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: October 14, 2003, 06:31:18 ; Search time 197.358 Seconds  
(without alignments)  
444.136 Million cell updates/sec  
Title: US-10-068-426-1  
Perfect score: 2906  
Sequence: 1 MPLLALLLLPSLPHPIC.....MHEALHHYTKSLSPGK 544

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 600653 seqs, 161128416 residues  
Total number of hits satisfying chosen parameters: 600653  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, AA:  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/FCR\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/FCRUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	544	15	US-10-068-426-1 Sequence 1, Appli
2	2894	99.6	544	15	US-10-068-426-2 Sequence 2, Appli
3	2881	99.1	544	15	US-10-068-426-3 Sequence 3, Appli
4	2825.5	97.2	531	15	US-10-068-426-4 Sequence 4, Appli
5	2819.5	97.0	531	15	US-10-068-426-6 Sequence 6, Appli
6	2812.5	96.8	531	15	US-10-068-426-5 Sequence 5, Appli
7	1598	55.0	302	15	US-10-068-426-7 Sequence 7, Appli
8	1586	54.6	302	15	US-10-068-426-8 Sequence 8, Appli
9	1555.5	53.5	301	15	US-10-068-426-9 Sequence 9, Appli
10	1539	53.0	290	15	US-10-068-426-10 Sequence 10, Appl
11	1533	52.8	290	15	US-10-068-426-12 Sequence 12, Appl
12	1526	52.5	290	15	US-10-068-426-11 Sequence 11, Appl
13	1236	42.5	313	12	US-09-825-580-4 Sequence 4, Appli
14	1236	42.5	313	15	US-10-211-786-4 Sequence 4, Appli
15	1232.5	42.4	388	9	US-09-784-623-16 Sequence 16, Appl

16	1232	42.4	580	11	US-09-972-268-36 Sequence 36, Appl
17	1230.5	42.3	453	9	US-09-802-077-8 Sequence 8, Appli
18	1230.5	42.3	453	9	US-09-802-096-8 Sequence 8, Appli
19	1230.5	42.3	453	11	US-09-925-179-8 Sequence 8, Appli
20	1229.5	42.3	451	9	US-09-920-171-14 Sequence 14, Appl
21	1229.5	42.3	451	9	US-09-920-171-16 Sequence 16, Appl
22	1229.5	42.3	451	9	US-09-920-171-18 Sequence 18, Appl
23	1229.5	42.3	451	11	US-09-925-179-65 Sequence 65, Appl
24	1229.5	42.3	451	11	US-09-925-179-66 Sequence 66, Appl
25	1229.5	42.3	451	12	US-10-113-996-14 Sequence 14, Appl
26	1229.5	42.3	451	12	US-10-113-996-16 Sequence 16, Appl
27	1229.5	42.3	451	12	US-10-113-996-18 Sequence 18, Appl
28	1229.5	42.3	451	12	US-10-292-869-2 Sequence 2, Appli
29	1229.5	42.3	451	12	US-09-792-938-2 Sequence 2, Appli
30	1225	42.2	694	10	US-09-935-868-18 Sequence 18, Appl
31	1225	42.2	694	12	US-10-282-162-18 Sequence 18, Appl
32	1225	42.2	694	15	US-10-287-035-18 Sequence 32, Appl
33	1225	42.2	793	10	US-09-935-868-32 Sequence 32, Appl
34	1225	42.2	793	12	US-10-282-162-32 Sequence 32, Appl
35	1225	42.2	793	15	US-10-287-035-32 Sequence 32, Appl
36	1224.5	42.1	977	12	US-10-357-653-1 Sequence 1, Appli
37	1224	42.1	401	9	US-09-859-361-9 Sequence 9, Appli
38	1223.5	42.1	451	11	US-09-925-179-68 Sequence 68, Appl
39	1223	42.1	382	15	US-10-207-655-307 Sequence 307, App
40	1223	42.1	680	7	US-08-469-583A-15 Sequence 15, Appl
41	1222	42.1	470	15	US-10-264-634-33 Sequence 33, Appl
42	1221.5	42.0	449	16	US-10-323-268-23 Sequence 23, Appl
43	1221	42.0	450	10	US-09-996-288-238 Sequence 238, App
44	1221	42.0	450	10	US-09-996-288-242 Sequence 242, App
45	1221	42.0	450	10	US-09-996-288-244 Sequence 244, App

ALIGNMENTS

RESULT 1  
US-10-068-426-1  
; Sequence 1, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: GP1b302-Ig  
US-10-068-426-1

Query Match 100.0%; Score 2906; DB 15; Length 544;  
Best Local Similarity 100.0%; Pred. No. 2.2e-224; Indels 0; Gaps 0;  
Matches 544; Conservative 0; Mismatches 0;  
QY 1 MPLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDTTILHSENLLY 60  
Db 1 MPLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDTTILHSENLLY 60  
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLFVLGTLDLSHNLQSLPLLGOTLPALT 120

```
Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPAITV 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
Qy 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENYVVMKQGVVDVKAMT:SNVASVQCDSNDKFPVYKPKGCPTLGDGSDTLDLYYPEEDT 300
Db 241 ENYVVMKQGVVDVKAMT:SNVASVQCDSNDKFPVYKPKGCPTLGDGSDTLDLYYPEEDT 300
Qy 301 EGDKVRATRVVVKPPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCVV 360
Db 301 EGDKVRATRVVVKPPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCVV 360
Qy 361 VDVSHEDEPKVFNWYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Db 361 VDVSHEDEPKVFNWYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Qy 421 SNKALPVIPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
Db 421 SNKALPVIPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
Qy 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSL 540
Db 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSL 540
Qy 541 SPKG 544
Db 541 SPKG 544
```

## RESULT 2

```
US-10-068-426-2
; Sequence 2, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPL302/2A-Ig
US-10-068-426-2
```

Query Match 99.6%; Score 2894; DB 15; Length 544;  
Best Local Similarity 99.6%; Pred. No. 2e-223;  
Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 MPLLALLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MPLLALLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
```

```
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPAITV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPAITV 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
Qy 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENYVVMKQGVVDVKAMT:SNVASVQCDSNDKFPVYKPKGCPTLGDGSDTLDLYYPEEDT 300
Db 241 ENYVVMKQGVVDVKAMT:SNVASVQCDSNDKFPVYKPKGCPTLGDGSDTLDLYYPEEDT 300
Qy 301 EGDKVRATRVVVKPPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCVV 360
Db 301 EGDKVAATATVVKPPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCVV 360
Qy 361 VDVSHEDEPKVFNWYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Db 361 VDVSHEDEPKVFNWYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Qy 421 SNKALPVIPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
Db 421 SNKALPVIPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
Qy 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSL 540
Db 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSL 540
Qy 541 SPKG 544
Db 541 SPKG 544
```

## RESULT 3

```
US-10-068-426-3
; Sequence 3, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPL302/4X-Ig
US-10-068-426-3
```

Query Match 99.1%; Score 2881; DB 15; Length 544;  
Best Local Similarity 99.3%; Pred. No. 2.2e-222;  
Matches 540; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 MPLLALLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MPLLALLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
```

QY 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVLTGLDLSHNOQSLPLGQTLPALTV 120  
DB 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVLTGLDLSHNOQSLPLGQTLPALTV 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
QY 181 AGLLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELTYFRRLQDNA 240  
DB 181 AGLLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELTYFRRLQDNA 240  
QY 241 ENVYWKQGVVVKAMTSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDT 300  
DB 241 ENVYWKQGVVVKAMTSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDT 300  
QY 301 EGDKVAATRTVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
DB 301 EGDKVAATRTVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
QY 361 VDVSHPEDPVKFNWYDGVGVHNAKTPREEOYNTYRVVSVLTVLHODWLNKGYKCKV 420  
DB 361 VDVSHPEDPVKFNWYDGVGVHNAKTPREEOYNTYRVVSVLTVLHODWLNKGYKCKV 420  
QY 421 SNKALPVPPIEKTISKAGOPREPQVVTLPSPREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
DB 421 SNKALPVPPIEKTISKAGOPREPQVVTLPSPREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
QY 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
QY 541 SPKG 544  
DB 541 SPKG 544

RESULT 4  
US-10-068-426-4  
; Sequence 4, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: 'Platlet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: 'Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290-Ig  
US-10-068-426-4  
Query Match 97.2%; Score 2825.5; DB 15; Length 531;  
Best Local Similarity 97.6%; Pred. No. 5.9e-218;  
Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTTLHLSENLLY 60  
|||||

DB 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTTLHLSENLLY 60  
QY 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVLTGLDLSHNOQSLPLGQTLPALTV 120  
DB 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVLTGLDLSHNOQSLPLGQTLPALTV 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
QY 181 AGLLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELTYFRRLQDNA 240  
DB 181 AGLLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELTYFRRLQDNA 240  
QY 241 ENVYWKQGVVVKAMTSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDT 300  
DB 241 ENVYWKQGVVVKAMTSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDT 300  
QY 301 EGDKVAATRTVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
DB 301 EGDKVAATRTVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
QY 361 VDVSHPEDPVKFNWYDGVGVHNAKTPREEOYNTYRVVSVLTVLHODWLNKGYKCKV 420  
DB 361 VDVSHPEDPVKFNWYDGVGVHNAKTPREEOYNTYRVVSVLTVLHODWLNKGYKCKV 420  
QY 421 SNKALPVPPIEKTISKAGOPREPQVVTLPSPREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
DB 421 SNKALPVPPIEKTISKAGOPREPQVVTLPSPREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
QY 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
QY 541 SPKG 544  
DB 528 SPKG 531

RESULT 5  
US-10-068-426-6  
; Sequence 6, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: 'Platlet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: 'Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290/1A-Ig  
US-10-068-426-6  
Query Match 97.0%; Score 2819.5; DB 15; Length 531;  
Best Local Similarity 97.4%; Pred. No. 1.8e-217;  
Matches 530; Conservative 0; Mismatches 1; Indels 13; Gaps 1;  
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTTLHLSENLLY 60  
|||||

```
Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTILHLSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLGLGOTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLGLGOTLPALT 120
Qy 121 LDVSFNLRLTSLPGALRGELQELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTEL 180
Db 121 LDVSFNLRLTSLPGALRGELQELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTEL 180
Qy 181 AGLINGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLINGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENVYVWKQGVDMKAMTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDDTLDLYYPEEDT 300
Db 241 ENVYVWKQGVDMKAMTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDDTLDLYYPEEDT 300
Qy 301 EGDVTRATRVVVFPTKARPHTCPPCPAPALGAPSVFLFPKPKDTLMSRTPETVCV 360
Db 301 EGDV-----RPHTCPPCPAPALGAPSVFLFPKPKDTLMSRTPETVCV 360
Qy 361 VDVSHEDEPVKFNWYVDCGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKYCKV 420
Db 361 VDVSHEDEPVKFNWYVDCGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKYCKV 420
Qy 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVWES 480
Db 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVWES 480
Qy 481 NGQENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSL 540
Db 481 NGQENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSL 527
Qy 541 SPKG 544
Db 528 SPKG 531
```

## RESULT 6

```
US-10-068-426-5
; Sequence 5, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068.426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290/2V-Ig
US-10-068-426-5
```

Query Match 96.8%; Score 2812.5; DB 15; Length 531;  
Best Local Similarity 97.2%; Pred. No. 6.5e-217;  
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

```
Qy 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTILHLSENLLY 60
Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRLTALPPDLPKDTILHLSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLGLGOTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLGLGOTLPALT 120
Qy 121 LDVSFNLRLTSLPGALRGELQELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTEL 180
Db 121 LDVSFNLRLTSLPGALRGELQELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTEL 180
Qy 181 AGLINGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLINGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENVYVWKQGVDMKAMTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDDTLDLYYPEEDT 300
Db 241 ENVYVWKQGVDMKAMTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDDTLDLYYPEEDT 300
Qy 301 EGDVTRATRVVVFPTKARPHTCPPCPAPALGAPSVFLFPKPKDTLMSRTPETVCV 360
Db 301 EGDV-----RPHTCPPCPAPALGAPSVFLFPKPKDTLMSRTPETVCV 360
Qy 361 VDVSHEDEPVKFNWYVDCGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKYCKV 420
Db 361 VDVSHEDEPVKFNWYVDCGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKYCKV 420
Qy 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVWES 480
Db 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVWES 480
Qy 481 NGQENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSL 540
Db 481 NGQENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSL 527
Qy 541 SPKG 544
Db 528 SPKG 531
```

## RESULT 7

```
US-10-068-426-7
; Sequence 7, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068.426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GPIb302
US-10-068-426-7
```

Query Match 55.0%; Score 1598; DB 15; Length 302;  
Best Local Similarity 100.0%; Pred. No. 6.2e-120;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY	17	HPICEVSKVASHLEVNCDQRNLTALPPDLPKDPTTILHLSNNLYTSLATMPYTRLTOL	76
DB	1	HPICEVSKVASHLEVNCDQRNLTALPPDLPKDPTTILHLSNNLYTSLATMPYTRLTOL	60
QY	77	NLDRCELTKLQVDGTLPVLGTLDSLHNOQLSILPGQLTPALTVLDVSNRLTSLPLGAL	136
DB	61	NLDRCELTKLQVDGTLPVLGTLDSLHNOQLSILPGQLTPALTVLDVSNRLTSLPLGAL	120
QY	137	RGELGELQELYLKXGNELKTLPPGGLTPTPKLEKLSLANNLTLPAGLNGLENDTLTLLQ	196
DB	121	RGELGELQELYLKXGNELKTLPPGGLTPTPKLEKLSLANNLTLPAGLNGLENDTLTLLQ	180
QY	197	ENSLYTIKPGFFGSHLLPFAFLHGNPWLCNCEILYFRRWLQDNAENYVWKQGVDKAMT	256
DB	181	ENSLYTIKPGFFGSHLLPFAFLHGNPWLCNCEILYFRRWLQDNAENYVWKQGVDKAMT	240
QY	257	SNVASVQCNSDKFPVYKYPGKGCPTLGDGDTLDLYYPEEDTEGDKVRAITRVVKFPT	316
DB	241	SNVASVQCNSDKFPVYKYPGKGCPTLGDGDTLDLYYPEEDTEGDKVRAITRVVKFPT	300
QY	317	KA 318	
DB	301	KA 302	

```

RESULT 8
US-10-068-426-8
; Sequence 8, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

```

Query Match	54.6%	Score 1586;	DB 15;	Length 302;
Best Local Similarity	99.3%	Prod. No. 5.7e-119;		
Matches 300;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	17	HPICEVSKVASHLEVNCDKRNLTALPPOLPKDITILHLSENLLYTFSLATLMPYTRLTQ	76	
DB	1	HPICEVSKVASHLEVNCDKRNLTALPPOLPKDITILHLSENLLYTFSLATLMPYTRLTQ	60	
QY	77	NLDRCELTKLQVGDGTPLVGTLTDLISHNQLQSLPLLGQTLPALTVLDVSPNRLTSLPLCAL	136	
DB	51	NLDRCELTKLQVGDGTPLVGTLTDLISHNQLQSLPLLGQTLPALTVLDVSPNRLTSLPLCAL	120	
QY	137	RLGLBELQLYIKGNELKTLPPGLLTPPTPKLEKLSLANNNLTELPAGLLNGLENIDTLILQ	196	
DB	121	RLGLBELQLYIKGNELKTLPPGLLTPPTPKLEKLSLANNNLTELPAGLLNGLENIDTLILQ	180	
QY	197	ENSLYTTIKPGFFGSHLLPFAFLHGNPWLNCBELLYFRFRWLQDQNAENVVYKQGVGVKAMT	256	

```

Db      181  ENSLYTPKPGFSGHLLPFAFLHGNPWLNCNCEILYFRFRLQDNNANVTYVWKQGVYDVKAMT 240
Qy      257  SNVASQCDNSDKFPVYKYPGKGCPTILGDEGDTLDLYDYYPEEDTEGDKVRAIRTVVAKFPT 316
Db      241  SNVASQCDNSDKFPVYKYPGKGCPTILGDEGDTLDLYDYYPEEDTEGDKVAATAITVVKFPT 300
Qy      317  KA 318
Db      301  KA 302

RESULT 9
US-10-068-426-9
; Sequence 9, Application US/10068426
; Publication NO. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(301)
; OTHER INFORMATION: Gp1b/4X
; US-10-068-426-9

```

Query Match	53.5%	Score	1555.5	DB	15	Length	301
Best Local Similarity	98.3%	Pred. No.	1.6e-116				
Matches	297	Conservative	1	Mismatches	3	Indels	1
						Gaps	1
Qy	17	HPICEVSKVASHLEVNCCKRNLTALPDP	KDTHLHSENNLLYTFESL	ATLMPY	TR	L	76
Db							
	1	HPICEVSKVASHLEVNCCKRNLTALPDP	KDTHLHSENNLLYTFESL	ATLMPY	TR	L	60
Qy	77	NLDRCETLKLQVDCGTL	PVLGTDLSHNQLQSLPL	LGOTLPAL	TVLDVS	FNRLTSLPL	136
Db							
	61	NLDRCETLKLQVDCGTL	PVLGTDLSHNQLQSLPL	LGOTLPAL	TVLDVS	FNRLTSLPL	120
Qy	137	RGIGELQELYLKGNELKYL	TPGLLTPPKL	EKLSIANNL	TETL	PAGLLNGLEN	196
Db							
	121	RGIGELQELYLKGNELKYL	TPGLLTPPKL	EKLSIANNL	TETL	PAGLLNGLEN	180
Qy	197	ENSLYTTIPKGFPGSHLLP	FAFLHGNPMLCNCEILY	FRFRLQDN	ABNVV	KQGV	256
Db							
	181	ENSLYTTIPKGFPGSHLLP	FAFLHGNPMLCNCEILY	FRFRLQDN	ABNVV	KQGV	240
Qy	257	SNVASVQCDNSDKFPV	KYPKGGCPT	TLGDEG	TDLYDYP	PEEDTEG	316
Db							
	241	SNVASVQCDNSDKFPV	KYPKGGCPT	TLGDEG	TDLYDYP	PEEDTEG	299
Qy	317	KA	318				
Db							
	300	KA	301				

RESULT 10  
US-10-068-426-10  
; Sequence 10, Application US/10068426  
; Publication No. US20030091576A1

```

; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: PlatiLet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290
US-10-068-426-10

Query Match          53.0%; Score 1539; DB 15; Length 290;
Best Local Similarity 100.0%; Pred. No. 3 1e-115;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

QY 197 ENSLYTIPIKGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDNAENVYVWKQGVVKAMT 256
Db 181 ENSLYTIPIKGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDNAENVYVWKQGVVKAMT 240

QY 257 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 290

US-10-068-426-11

RESULT 11
US-10-068-426-12
; Sequence 12, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: PlatiLet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290/2V
US-10-068-426-11

Query Match          52.5%; Score 1526; DB 15; Length 290;
Best Local Similarity 99.3%; Pred. No. 3.4e-114;
Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196

```

```

; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gb1b290/1A
US-10-068-426-12

Query Match          52.8%; Score 1533; DB 15; Length 290;
Best Local Similarity 99.7%; Pred. No. 9.5e-115;
Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

QY 197 ENSLYTIPIKGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDNAENVYVWKQGVVKAMT 256
Db 181 ENSLYTIPIKGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDNAENVYVWKQGVVKAMT 240

QY 257 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 290

US-10-068-426-11

RESULT 12
US-10-068-426-11
; Sequence 11, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: PlatiLet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gb1b290/2V
US-10-068-426-11

Query Match          52.5%; Score 1526; DB 15; Length 290;
Best Local Similarity 99.3%; Pred. No. 3.4e-114;
Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196

```

Db 121 RGLGELQELYLKGNELKTLPPGLLTTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFPGSHLLPFAFLHGNPWLNCNCEILYPRRWLQDAENNVYWKQGVVKAMT 256  
Db 181 ENSLYTIPIKGFPGSHLLPFAFLHGNPWLNCNCEILYPRRWLQDAENNVYWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPYKYPKPGKCPDGLDGDGDTLDYDYPBEDTEGDKVR 306  
Db 241 SNVASVQCDNSDKFPYKYPKPGKCPDGLDGDGDTLDYDYPBEDTEGDKVR 290

## RESULT 13

US-09-825-580-4  
; Sequence 4, Application US/09825580  
; Publication No. US20030166521A1  
; GENERAL INFORMATION:  
; APPLICANT: Eppihimer, Michael J.  
; APPLICANT: Schaub, Robert G.  
; APPLICANT: Harris, Alan  
; TITLE OF INVENTION: Inhibition of Thrombosis by Treatment with  
; P-Selectin Antagonists  
; FILE REFERENCE: GPN-5398  
; CURRENT APPLICATION NUMBER: US/09/825,580  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/193,787  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-825-580-4

Query Match 42.5%; Score 1236; DB 12; Length 313;  
Best Local Similarity 88.6%; Pred. No. 6.7e-91;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;  
QY 292 YDYYPEDETEGDKVRATRTV-----VKPPTKARPHTCPPCPAPEALGAPSVFLFP 341  
Db 51 YDFLPETEPPEMLRNSTDTTFLPGTGPSTTVEPAARPHTCPPCPAPEALGAPSVFLFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
Db 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 170  
QY 402 VLTVLHODWLNGKEYKCKVSNKALPVPPIETISKAKGQPREPOVYTLPPSREEMTKNQVS 461  
Db 171 VLTVLHODWLNGKEYKCKVSNKALPVPPIETISKAKGQPREPOVYTLPPSREEMTKNQVS 230  
QY 462 LTCLVKGFPSPDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 521  
Db 231 LTCLVKGFPSPDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 290  
QY 522 CSVMHEALHNHYTOKLSLSLSPGK 544  
Db 291 CSVMHEALHNHYTOKLSLSLSPGK 313

## RESULT 14

US-10-211-786-4  
; Sequence 4, Application US/10211786  
; Publication No. US20030083258A1  
; GENERAL INFORMATION:  
; APPLICANT: Michael J. Eppihimer  
; APPLICANT: Robert G. Schaub  
; APPLICANT: Ronald Tuma  
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM  
; FILE REFERENCE: 8702.0099-00000  
; CURRENT APPLICATION NUMBER: US/10/211,786  
; CURRENT FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: 60/309,816

; PRIOR FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-211-786-4

Query Match 42.5%; Score 1236; DB 15; Length 313;  
Best Local Similarity 88.6%; Pred. No. 6.7e-91;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;

QY 292 YDYYPEDETEGDKVRATRTV-----VKPPTKARPHTCPPCPAPEALGAPSVFLFP 341  
Db 51 YDFLPETEPPEMLRNSTDTTFLPGTGPSTTVEPAARPHTCPPCPAPEALGAPSVFLFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
Db 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 170  
QY 402 VLTVLHODWLNGKEYKCKVSNKALPVPPIETISKAKGQPREPOVYTLPPSREEMTKNQVS 461  
Db 171 VLTVLHODWLNGKEYKCKVSNKALPVPPIETISKAKGQPREPOVYTLPPSREEMTKNQVS 230  
QY 462 LTCLVKGFPSPDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 521  
Db 231 LTCLVKGFPSPDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 290  
QY 522 CSVMHEALHNHYTOKLSLSLSPGK 544  
Db 291 CSVMHEALHNHYTOKLSLSLSPGK 313

## RESULT 15

US-09-784-623-16  
; Sequence 16, Application US/09784623  
; Patent No. US20020009454A1  
; GENERAL INFORMATION:  
; APPLICANT: Boone, Thomas C.  
; APPLICANT: Hershenson, Susan  
; APPLICANT: Bevilacqua, Michael P.  
; APPLICANT: Collins, David S.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
; TITLE OF INVENTION: DISEASES  
; FILE REFERENCE: A-365F  
; CURRENT APPLICATION NUMBER: US/09/784,623  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: 09/131,247  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: PCT/US 97/02131  
; PRIOR FILING DATE: 1997-02-10  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Human  
US-09-784-623-16

Query Match 42.4%; Score 1232.5; DB 9; Length 388;  
Best Local Similarity 63.1%; Pred. No. 1.7e-90;  
Matches 250; Conservative 21; Mismatches 70; Indels 55; Gaps 6;

QY 171 LANNLTLPAGLNG-----LENLDTLLIQNSLYTIPIKGFPGSHLLPFAFLHGNPWL 225  
Db 26 LRNN---QLVAGYLQGGNVNLEBKIDVVPTEPHAL-----FLGIHGGRMCL 68  
QY 226 NCEILYFRWLQDAENNVYWKQGVVKAMTSNVASVQCDNSDKFPYKYPKPGKCP--- 281  
Db 69 SCVKSGDETRQLQLEAVNI-----TDLSSNRKODKRFATIRSDSGPTTSFSAACPNFL 122  
QY 282 -----TLGDBGDTLDYDYPBEDTEGDKVRATRTVVKPPTKARPHTCPPCPA 328

```

Db 123 CTAMEADQPVSILTNPDEGVMTKFPYQDEDEAAEPKSSDKT-----HTCPCPA 172
Qy 329 PEALGAPSVLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 388
Db 173 PELIGGPSVLFPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 232
Qy 389 REEOYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVPPIETISKAKGQPREPOVYTL 446
Db 233 REEOYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVPPIETISKAKGQPREPOVYTL 292
Qy 449 PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT 508
Db 293 PPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT 352
Qy 509 VDKSRWQOGNVPFSCVMHEALHNHYTOKSLSLSPGK 544
Db 353 VDKSRWQOGNVPFSCVMHEALHNHYTOKSLSLSPGK 388

```

Search completed: October 14, 2003, 06:48:21  
Job time : 199.358 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: October 14, 2003, 06:17:26 ; Search time 22.266 Seconds  
(without alignments)  
1033.731 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLILLLLPLPHPIC.....MHEALHNHYTKLSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents AA:\*  
1: /cgn2\_6/protdata/2/aa/5A\_COMB.pep:\*  
2: /cgn2\_6/protdata/2/aa/5B\_COMB.pep:\*  
3: /cgn2\_6/protdata/2/aa/6A\_COMB.pep:\*  
4: /cgn2\_6/protdata/2/aa/6B\_COMB.pep:\*  
5: /cgn2\_6/protdata/2/aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/protdata/2/aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1606	55.3	320	1 US-07-613-083B-1	Sequence 1, Appli
2	1599	55.0	610	1 US-07-821-717B-6	Sequence 6, Appli
3	1599	55.0	610	1 US-08-119-262B-6	Sequence 6, Appli
4	1599	55.0	610	1 US-08-135-929A-11	Sequence 11, Appli
5	1599	55.0	610	1 US-08-234-265A-11	Sequence 11, Appli
6	1236	42.5	313	3 US-08-713-558F-36	Sequence 36, Appli
7	1232.5	42.4	388	3 US-09-131-247-16	Sequence 16, Appli
8	1230.5	42.3	453	3 US-08-466-151-8	Sequence 8, Appli
9	1230.5	42.3	453	4 US-08-466-163B-8	Sequence 8, Appli
10	1229.5	42.3	451	2 US-08-887-352B-14	Sequence 14, Appli
11	1229.5	42.3	451	2 US-08-887-352B-16	Sequence 16, Appli
12	1229.5	42.3	451	2 US-08-887-352B-18	Sequence 18, Appli
13	1229.5	42.3	451	3 US-08-466-151-65	Sequence 65, Appli
14	1229.5	42.3	451	3 US-09-109-207C-14	Sequence 14, Appli
15	1229.5	42.3	451	3 US-09-109-207C-16	Sequence 16, Appli
16	1229.5	42.3	451	3 US-09-109-207C-18	Sequence 18, Appli
17	1229.5	42.3	451	3 US-09-282-505-2	Sequence 2, Appli
18	1229.5	42.3	451	3 US-09-054-255-2	Sequence 2, Appli
19	1229.5	42.3	451	3 US-09-296-005-14	Sequence 14, Appli
20	1229.5	42.3	451	3 US-09-296-005-16	Sequence 16, Appli
21	1229.5	42.3	451	3 US-09-296-005-18	Sequence 18, Appli
22	1229.5	42.3	451	4 US-09-282-846-2	Sequence 2, Appli
23	1229.5	42.3	451	4 US-09-680-145-2	Sequence 2, Appli
24	1225	42.2	694	4 US-09-313-942-18	Sequence 18, Appli
25	1225	42.2	793	4 US-09-313-942-32	Sequence 32, Appli
26	1224.5	42.1	977	4 US-09-590-656-1	Sequence 1, Appli
27	1224.5	42.1	977	4 US-09-733-764-1	Sequence 1, Appli

Sequence 15, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 30, Appli  
Sequence 43, Appli  
Sequence 43, Appli  
Sequence 43, Appli  
Sequence 43, Appli  
Sequence 8, Appli  
Sequence 43, Appli  
Sequence 71, Appli  
Sequence 71, Appli  
Sequence 71, Appli  
Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-07-613-083B-1  
; Sequence 1, Application US/07613083B  
; Patent No. 5340727  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Ware, Jerry, inventors  
; APPLICANT: on behalf of Scripps Clinic and Research  
; APPLICANT: Foundation  
; TITLE OF INVENTION: GPIb Fragments and Recombinant  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scripps Clinic and Research  
; ADDRESSEE: Foundation  
; STREET: 10666 No. 5340727th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb  
; COMPUTER: AST Bravo IBM PC comp. (386SX)  
; OPERATING SYSTEM: MS DOS version 3.2  
; SOFTWARE: WordPerfect 5.1 conv. to ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/613,083B  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: This appl. is a c-i-p of  
; APPLICATION NUMBER: U.S. 07/470,674  
; FILING DATE: 04-Jan-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barron, Alexis  
; REGISTRATION NUMBER: 22,702  
; REFERENCE/DOCKET NUMBER: P16,569-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 923-4466  
; TELEFAX: (215) 923-2189  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320  
; TYPE: AMINO ACID  
; STRANDEDNESS: No. 5340727 applicable  
; TOPOLOGY: Linear  
US-07-613-083B-1

Query Match 55.3%; Score 1606; DB 1; Length 320;  
Best Local Similarity 99.0%; Pred. No. 1.5e-132;  
Matches 305; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 17 HPICEVSKVASHLEVNCNLTALPPDLPKDPTILHLSENLLYTFSLATLMPYTRLTQL 76  
Db 1 HPICEVSKVASHLEVNCNLTALPPDLPKDPTILHLSENLLYTFSLATLMPYTRLTQL 60  
QY 77 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGLQTLPALTVLDVSNRLTSLPLGAL 136  
Db 61 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGLQTLPALTVLDVSNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKLTLPGLLTTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
Db 121 RGLGELQELYLKGNELKLTLPGLLTTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNLTLYFRRLQNAENVYVWKQGVVKAMT 256  
Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNLTLYFRRLQNAENVYVWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKPGKCPGLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 316  
Db 241 SNVASVQCDNSDKFPVYKPGKCPGLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 300  
QY 317 KARPHTCP 324  
Db 301 KA--HTTP 306

## RESULT 2

US-07-821-717B-6  
; Sequence 6, Application US/07821717B  
; Patent No. 5298239  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/821,717B  
; FILING DATE: 15-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.

; AUTHORS: Papayannopoulou, Thalia  
; AUTHORS: Roth, Gerald J.  
; TITLE: Cloning of the alpha chain of human  
; TITLE: Platelet glycoprotein Ib: A transmembrane protein with homology  
; TITLE: to leucine-rich alpha-2-glycoprotein  
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
; VOLUME: 84  
; PAGES: 5615-5619  
; DATE: AUG-1987  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
; PUBLICATION INFORMATION:  
; AUTHORS: Zimmerman, Theodore S.  
; AUTHORS: Ruggeri, Zaverio M.  
; AUTHORS: Houghten, Richard A.  
; AUTHORS: Vincete, Vincete  
; AUTHORS: Mohri, Hiroshi  
; TITLE: Proteolytic fragments and synthetic  
; TITLE: peptides that block the binding of von Willebrand factor to the  
; TITLE: platelet membrane glycoprotein Ib  
; DOCUMENT NUMBER: EP 0 317 278 A2  
; FILING DATE: 16-NOV-1988  
; PUBLICATION DATE: 24-MAY-1989  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
; US-07-821-717B-6

Query Match 55.0%; Score 1599; DB 1; Length 610;  
Best Local Similarity 86.3%; Pred. No. 1.6e-131;  
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCNLTALPPDLPKDPTILHLSENLLYTFSLATLMPYTRLTQL 76  
Db 1 HPICEVSKVASHLEVNCNLTALPPDLPKDPTILHLSENLLYTFSLATLMPYTRLTQL 60  
QY 77 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGLQTLPALTVLDVSNRLTSLPLGAL 136  
Db 61 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGLQTLPALTVLDVSNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKLTLPGLLTTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
Db 121 RGLGELQELYLKGNELKLTLPGLLTTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNLTLYFRRLQNAENVYVWKQGVVKAMT 256  
Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNLTLYFRRLQNAENVYVWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKPGKCPGLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 316  
Db 241 SNVASVQCDNSDKFPVYKPGKCPGLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 300  
QY 317 KARPHTCP-----PCPA---PEALGAPSVFLPPPK--PKDTL-----MISR 352  
Db 301 KA--HTTPWGLFYSMSTASLDSQMPSSLHPTQBSTKEQTTFPPRWTNFTLHMBESITFSK 358  
QY 353 TPEVT 357  
Db 359 TPKST 363

## RESULT 3

US-08-119-262B-6  
; Sequence 6, Application US/08119262B  
; Patent No. 5492809  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051

CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,462B  
FILING DATE: 09-SEP-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/821,717  
FILING DATE: 15-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/22  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Lopez, Jose A.  
AUTHORS: Chung, Dominic W.  
AUTHORS: Fujikawa, Kazuo  
AUTHORS: Hagen, Frederick S.  
AUTHORS: Papayannopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human platelet glycoprotein Ib: A transmembrane protein  
TITLE: leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic peptides that block the binding of von Willebrand  
TITLE: membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-08-119-262B-6  
Query Match 55.0%; Score 1599; DB 1; Length 610;  
Best Local Similarity 86.3%; Pred. No. 1.6e-131;  
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTSFSLATLMPYTRLTOL 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTSFSLATLMPYTRLTOL 60  
QY 77 NLDRCETLKQVDGTLFVLGTLDSLHNOQLSPLLGQTLPALTVLDVSNELTSLPLGAL 136  
Db 61 NLDRCETLKQVDGTLFVLGTLDSLHNOQLSPLLGQTLPALTVLDVSNELTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKLTTPGGLTTPPKLEKLSLANNLTLPVGLLNGLENLDTLLQ 196

121 RGLGELQELYLKGNELKLTTPGGLTTPPKLEKLSLANNLTLPVGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCBILYFRWLQDNAENVYWKQGVYKAMT 256  
Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCBILYFRWLQDNAENVYWKQGVYKAMT 240  
QY 257 SNVSVQCDNSDKFPVYKYPGKGCPTLGGEGDLDLYYYPEDETEGDKVRATRTVVKFPT 316  
Db 241 SNVSVQCDNSDKFPVYKYPGKGCPTLGGEGDLDLYYYPEDETEGDKVRATRTVVKFPT 300  
QY 317 KARPHTCP-----PCPA---PEALGAPSVFLPPK--PKDTL-----MISR 352  
Db 301 KA--HTTPMGLFYSWASTASLDQMPSSLHPTQBSTKEQITTFPRWTPNFTLHMSITFSK 358  
QY 353 TPEVT 357  
Db 359 TPKST 363  
RESULT 4  
US-08-135-929A-11  
; Sequence 11, Application US/08135929A  
; Patent No. 5593959  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; APPLICANT: Pincus, Matthew R.  
; TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha  
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/135,929A  
; FILING DATE: 14-OCT-1993  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/23  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; TELEX: 978450  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-135-929A-11  
Query Match 55.0%; Score 1599; DB 1; Length 610;  
Best Local Similarity 86.3%; Pred. No. 1.6e-131;  
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTSFSLATLMPYTRLTOL 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTSFSLATLMPYTRLTOL 60

with homology

factor to the

QY 77 NLDRCETLKLQVDGTLPVLTGTLDSHNOLOSLPLIGOTLPALTVDVSNFRLTSLPLGAL 136  
DB 61 NLDRCETLKLQVDGTLPVLTGTLDSHNOLOSLPLIGOTLPALTVDVSNFRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQGVVKAMT 256  
DB 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPKGCPTLGDGDTLDYDYPEDTEGDKVRATRTVVKPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPKGCPTLGDGDTLDYDYPEDTEGDKVRATRTVVKPT 300  
QY 317 KARPHTCP-----PCPA---PEALGAFSVF--PKDTL-----MISR 352  
DB 301 KA--HTTPWGLFYSWSTASLDSQMPSSLHPTQESTKEQTTFPRTWPNFTLHMESITFSK 358  
QY 353 TPEVT 357  
DB 359 TPKST 363

## RESULT 5

US-08-234-265A-11  
; Sequence 11, Application US/08234265A  
; Patent No. 5624817  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; APPLICANT: Pincus, Matthew R.  
; TITLE OF INVENTION: Mutations in the Gene Encoding the  
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/234,265A  
; FILING DATE: 28-APR-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/24  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; TELEX: 978450  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-234-265A-11

Query Match 55.0%; Score 1599; DB 1; Length 610;  
Best Local Similarity 86.3%; Pred. No. 1.6e-131;  
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;

QY 17 HPICEVSKVASHLEVNCCKNLTALPPDLPKDTTILHLSENLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSKVASHLEVNCCKNLTALPPDLPKDTTILHLSENLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCETLKLQVDGTLPVLTGTLDSHNOLOSLPLIGOTLPALTVDVSNFRLTSLPLGAL 136  
DB 61 NLDRCETLKLQVDGTLPVLTGTLDSHNOLOSLPLIGOTLPALTVDVSNFRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQGVVKAMT 256  
DB 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPKGCPTLGDGDTLDYDYPEDTEGDKVRATRTVVKPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPKGCPTLGDGDTLDYDYPEDTEGDKVRATRTVVKPT 300  
QY 317 KARPHTCP-----PCPA---PEALGAFSVF--PKDTL-----MISR 352  
DB 301 KA--HTTPWGLFYSWSTASLDSQMPSSLHPTQESTKEQTTFPRTWPNFTLHMESITFSK 358  
QY 353 TPEVT 357  
DB 359 TPKST 363

RESULT 6  
US-08-713-556F-36  
; Sequence 36, Application US/08713556F  
; Patent No. 6277975  
; GENERAL INFORMATION:  
; APPLICANT: Larsen, Glenn  
; APPLICANT: Sako, Dianne  
; APPLICANT: Chang, Xiao Jia  
; APPLICANT: Veldman, Geertuida M.  
; APPLICANT: Cumming, Dale  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Shaw, Gray  
; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LEGAL AFFAIRS  
; STREET: 87 CAMBRIDGE PARK DRIVE  
; CITY: CAMBRIDGE  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/713,556F  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/965,662  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/112,608  
; FILING DATE: 26-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/10168  
; FILING DATE: 22-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/235,398  
; FILING DATE: 28-APR-1994  
; PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 08/316,305  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/428,734  
FILING DATE: 25-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWN, SCOTT A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI 5213P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-713-556F-36

Query Match 42.5%; Score 1236; DB 3; Length 313;  
Best Local Similarity 88.6%; Pred. No. 3.6e-100;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;  
QY 292 YDYYPEEDTEGDKVRAETV-----VKFPTKARPHTCPPCAPEALGAPSVLEFP 341  
DB 51 YDFLPETEPPEMLRNSTDTPLTGTPPESTTVEPARPHTCPPCAPEALGAPSVLEFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 401  
DB 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 170  
QY 402 VLTVLHODWLNKGYCKVSKNKPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQV 461  
DB 171 VLTVLHODWLNKGYCKVSKNKPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQV 230  
QY 462 LTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSLKLTVDKSRWQGNVFS 521  
DB 231 LTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSLKLTVDKSRWQGNVFS 290  
QY 522 CSVMHEALHNHYTQKSLSLSPCK 544  
DB 291 CSVMHEALHNHYTQKSLSLSPCK 313

RESULT 7  
US-09-131-247-16  
Sequence 16, Application US/09131247  
Patent No. 6294170  
GENERAL INFORMATION:  
APPLICANT: Boone, Thomas C.  
APPLICANT: Herhenson, Susan  
APPLICANT: Bevilacqua, Michael P.  
APPLICANT: Collins, David S.  
TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
TITLE OF INVENTION: DISEASES  
FILE REFERENCE: A-365P  
CURRENT APPLICATION NUMBER: US/09/131,247  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: 60/055,185  
EARLIER FILING DATE: 1997-08-08  
EARLIER APPLICATION NUMBER: PCT/US 97/02131  
EARLIER FILING DATE: 1997-02-10  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Human  
US-09-131-247-16  
Query Match 42.4%; Score 1232.5; DB 3; Length 388;

Best Local Similarity 63.1%; Pred. No. 1e-99;  
Matches 250; Conservative 21; Mismatches 70; Indels 55; Gaps 6;  
QY 171 LANNLTLPAGLNG-----LENLDTLLQENSLYTIPTKGFSGHLLPFAFLHGNPWL 225  
DB 26 LANN---QLVAGYLQGPNNVLEEKIDVVPPIEPHAL-----FLGIHGKMKCL 68  
QY 226 NCEILYFRWLQDANENYVWKQGVDKAMTSNVASVQCDNSDKFPVYKYPKGCPCP----- 281  
DB 69 SCVKSGETRLQLEAVNL-----TDLSENKQDKRFAFIRSDSGPTTSFESAACPGWFL 122  
QY 282 -----TLGDEGDTLDYDYPEEDTEGDKVRAETVTVVKFPTKARPHTCPPCPA 328  
DB 123 CTAMEADQPVSLTNMDEGVMVTKFYFQDEAAAEPKSSDKT-----HTCPCCPA 172  
QY 329 PEALGAPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 388  
DB 173 PELLGGPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 232  
QY 389 REEQYNSTYRVVSVLTVLHODWLNKGYCKVSKNKPVPPIEKTISKAKGQPREPOVYTL 448  
DB 233 REEQYNSTYRVVSVLTVLHODWLNKGYCKVSKNKPVPPIEKTISKAKGQPREPOVYTL 292  
QY 449 PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSLKLT 508  
DB 293 PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSLKLT 352  
QY 509 VDKSRWQGNVFS CSVMHEALHNHYTQKSLSLSPCK 544  
DB 353 VDKSRWQGNVFS CSVMHEALHNHYTQKSLSLSPCK 388

RESULT 8  
US-08-466-151-8  
Sequence 8, Application US/08466151  
Patent No. 6037453  
GENERAL INFORMATION:  
APPLICANT: Jardieu, Paula M.  
APPLICANT: Presta, Leonard G.  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,151  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/466163  
FILING DATE: 06-Jun-1995  
APPLICATION NUMBER: 08/405617  
FILING DATE: 15-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/185899  
FILING DATE: 26-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/879495  
FILING DATE: 07-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/744768  
FILING DATE: 14-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.



Best Local Similarity 61.7%; Pred. No. 2.3e-99;		
Matches 254;	Conservative 16;	Mismatches 53; Indels 89; Gaps 8;
QY	188	ENLDTLLQENSLVTPKGF-----GSHLLPFLAFLHGNPWLGCNCEILYFRRLQDNAENV 243
Db	74	DSKNTFYLQNSLRAEATAVYCARGSH-----YFGHW-----HF 108
QY	244	YWKQGVDDKAMTSNVAQCDNSDKFPV---YKYPKGCGPTLGDGDTDLDYYPEEDT 300
Db	109	AVWGQGTIV-----TVSSASTKGPSVFLAPSSKSTSGGTAALG---CLVKDYFPPEVT 159
QY	301	-----EGDKVRASTRVVKFPTKA----- 318
Db	160	VSMNKGALTSGVHTFPVLQSSGLYSLSVVTPVPSLSLGTQTYICNVNHHKPSNTKVDKKV 219
QY	319	-----RPHCTPCCPAPEALCAPSVLFPKPKDPTLMSRTPEVTCVVDVSHEDPEVKF 372
Db	220	EPKSCDKHTCTPCCPAPELLGGPSVFLFPKPKDPTLMSRTPEVTCVVDVSHEDPEVKF 279
QY	373	NWTVDGVVEVHNAKTKPREEQVNSYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKT 432
Db	280	NWTVDGVVEVHNAKTKPREEQVNSYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
QY	433	ISKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 492
Db	340	ISKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 399
QY	493	PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 544
Db	400	PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 451

```

RESULT 11
US-08-887-352B-16
; Sequence 16, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; ;
; TITLE OF INVENTION: Improving Polypeptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,352B
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-887-352B-16

```

Matches	254;	Conservative	16;	Mismatches	53;	Indels	89;	Gaps	8;
---------	------	--------------	-----	------------	-----	--------	-----	------	----

  

Qy	188	ENLDTLLQLQNSLYTIPKGF	-----GSHLLPFAFLHGNPMLNCCEILYFRRLWQDQNAENV	243
Db	74	DSKNTPYLQMSLRAEDTAVVYCARGSH	-----YFGHW-----	108
Qy	244	YVWQGVVDVKAMTSNVASVQCDNSDKPV	-----YKYPGKGCTFLGDEGDTLDLYDYPEEDT	300
Db	109	AVWGQGTLLV	-----TVSSASTKGPFSVFLPAPSSKTSQGTAAALG	159
Qy	301	-----EGDKVRAETVVKPETKA	-----	318
Db	160	VSWNGSALTSVGHITPAVLQSSGLSVTVTFSSSLGTQTYICNVNHPKSTNVDKVV	219	
Qy	319	-----RPHCTPCCPAPBALGAPSVFLFPFKPKDQTLMI	SRTPEVTCVVVDVSHEDPEVKF	372
Db	220	EPKSCDKTHCTPCCPAPBLGSPSVFLFPFKPKDQTLMI	SRTPEVTCVVVDVSHEDPEVKF	279
Qy	373	NYVVGVEVHNAKTPRBEQVNSTYRVVSVLTVLHODWLNGKEYCKVSNKALPVP	IEKT	432
Db	280	NYVVGVEVHNAKTPRBEQVNSTYRVVSVLTVLHODWLNGKEYCKVSNKALP	IEKT	339
Qy	433	ISAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOGPENNYKTP	492	
Db	340	ISAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOGPENNYKTP	399	
Qy	493	PVLDSGSPFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPK	544	
Db	400	PVLDSGSPFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPK	451	

  

RESULT 12

US-08-887-352B-18

; Sequence 18, Application US/08887352B

; Patent No. 5994511

; GENERAL INFORMATION:

; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe

; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of

; TITLE OF INVENTION: Improving Polypeptides

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/887,352B

; FILING DATE: 03-Jul-1997

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Svoboda, Craig G.

; REGISTRATION NUMBER: 39,044

; REFERENCE/DOCKET NUMBER: P1123

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-1489

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 451 amino acids

; TYPE: Amino Acid

; TOPOLOGY: Linear

US-08-887-352B-18

  

Query Match	42.3%;	Score	1229.5;	DB 2;	Length	451;			
Best Local Similarity	61.7%;	Pred. No.	2.3e-99;						
Matches	254;	Conservative	16;	Mismatches	53;	Indels	89;	Gaps	8;

```

QY 188 ENLDTLLQNSLYTTPKGF-----GSHLLPFAFLHGNPWLNCNCEILYFRWLQDAENV 243
Db 74 DSKNTFYLQNSLRADTAIVYTCARGSH-----YFGHW-----HF 108
QY 244 YVMKQGVVVKAMTSNVASVQCONSDKFPV---YKYPGKGCPTLGDGDTDLVYYPEEDT 300
Db 109 AVWGQGLV-----TVSSASTKGPSVFPPLAPSKSTSGGTAALG-----CLVKDYFPEPVT 159
QY 301 -----EGDKVRATRTVVKFPPTKA----- 318
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSVTVVPSSSLGTQTYICNVNHPKSNKALPVPIEKT 219
QY 319 -----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 279
QY 373 NMYVDGVEVHNATKPREQYNSYTRVSVLTVLHODWLNKGYCKVSNKALPVPIEKT 432
Db 280 NMYVDGVEVHNATKPREQYNSYTRVSVLTVLHODWLNKGYCKVSNKALPVPIEKT 339
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 399
QY 493 PVLDSGSEFELYSLKLTVDKSRMOQGNVFCSSVMHEALHNHYTKSLSLSPGK 544
Db 400 PVLDSGSEFELYSLKLTVDKSRMOQGNVFCSSVMHEALHNHYTKSLSLSPGK 451

RESULT 13
US-08-466-151-65
; Sequence 65, Application US/08466151
; Patent No. 6037453
; GENERAL INFORMATION:
; APPLICANT: Presta, Paula M.
; APPLICANT: Jardieu, Leonard G.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,151
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466163
; FILING DATE: 06-Jun-1995
; APPLICATION NUMBER: 08/405617
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/185899
; FILING DATE: 26-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/879495
; FILING DATE: 07-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P2C1D1

```

## TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-466-151-65

```

```

Query Match 42.3%; Score 1229.5; DB 3; Length 451;
Best Local Similarity 61.7%; Pred. No. 2.3e-99;
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

QY 188 ENLDTLLQNSLYTTPKGF-----GSHLLPFAFLHGNPWLNCNCEILYFRWLQDAENV 243
Db 74 DSKNTFYLQNSLRADTAIVYTCARGSH-----YFGHW-----HF 108
QY 244 YVMKQGVVVKAMTSNVASVQCONSDKFPV---YKYPGKGCPTLGDGDTDLVYYPEEDT 300
Db 109 AVWGQGLV-----TVSSASTKGPSVFPPLAPSKSTSGGTAALG-----CLVKDYFPEPVT 159
QY 301 -----EGDKVRATRTVVKFPPTKA----- 318
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSVTVVPSSSLGTQTYICNVNHPKSNKALPVPIEKT 219
QY 319 -----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 279
QY 373 NMYVDGVEVHNATKPREQYNSYTRVSVLTVLHODWLNKGYCKVSNKALPVPIEKT 432
Db 280 NMYVDGVEVHNATKPREQYNSYTRVSVLTVLHODWLNKGYCKVSNKALPVPIEKT 339
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 399
QY 493 PVLDSGSEFELYSLKLTVDKSRMOQGNVFCSSVMHEALHNHYTKSLSLSPGK 544
Db 400 PVLDSGSEFELYSLKLTVDKSRMOQGNVFCSSVMHEALHNHYTKSLSLSPGK 451

```

## RESULT 14

```

US-09-109-207C-14
; Sequence 14, Application US/09109207C
; Patent No. 6172213
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept
; FILE REFERENCE: P1123R1
; CURRENT APPLICATION NUMBER: US/09/109,207C
; CURRENT FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/051,554
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 14
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: Artificial
; LOCATION: 1-451
; OTHER INFORMATION: Heavy chain sequence derived from MAE11
; US-09-109-207C-14

```

```

Query Match 42.3%; Score 1229.5; DB 3; Length 451;
Best Local Similarity 61.7%; Pred. No. 2.3e-99;
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

QY 188 ENLDTLLQNSLYTTPKGF-----GSHLLPFAFLHGNPWLNCNCEILYFRWLQDAENV 243
Db 74 DSKNTFYLQNSLRADTAIVYTCARGSH-----YFGHW-----HF 108

```

```
QY 244 YVMKQGVDMKAMTSNVASVQDNDKFPV---YKYPGKGCPTLGDGDTLDLYDYPEEDT 300
Db 109 AVWQOGLV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG-----CLVKDYFPEPVT 159
QY 301 -----EGDKVRAIRTVVVKFPTKA-----
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKYDKKV 219
QY 319 -----RPHTCCPCPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCCPCPAPELGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 279
QY 373 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
Db 280 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTT 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTT 399
QY 493 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 544
Db 400 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 451
```

## RESULT 15

```
US-09-109-207C-16
; Sequence 16, Application US/09109207C
; Patent No. 6172213
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypeptide
; FILE REFERENCE: P1123R1
; CURRENT APPLICATION NUMBER: US/09/109,207C
; CURRENT FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/051,554
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 16
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: Artificial
; LOCATION: 1-451
; OTHER INFORMATION: Heavy chain sequence derived from MAEL1
US-09-109-207C-16
```

```
Query Match 42.3%; Score 1229.5; DB 3; Length 451;
Best Local Similarity 61.7%; Pred. No. 2.3e-99;
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

QY 198 ENLDPTLLQENSLYTIPIKGF---GSHLLPFAFLHGNPMLCNCEILYFRWLQDNAENV 243
Db 74 DSKNTFYLMNSLRAEDTAVYYCARGSH-----YFGHW-----HF 108
QY 244 YVMKQGVDMKAMTSNVASVQDNDKFPV---YKYPGKGCPTLGDGDTLDLYDYPEEDT 300
Db 109 AVWQOGLV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG-----CLVKDYFPEPVT 159
QY 301 -----EGDKVRAIRTVVVKFPTKA-----
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKYDKKV 219
QY 319 -----RPHTCCPCPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCCPCPAPELGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 279
QY 373 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
Db 280 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
```

```
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTT 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTT 399
QY 493 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 544
Db 400 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 451
```

Search completed: October 14, 2003, 06:32:09  
Job time : 26.266 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 24.6977 Seconds  
(without alignments)  
2067.627 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPELLLLLLLPSPLPHPTIC.....MHEALHNYTQKSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1615.5	56.9	626	1 NBHUIA	platelet glycoprot
2	1205	42.4	330	1 GHU	Ig gamma-1 chain C
3	1203	42.4	374	2 S69339	Ig heavy chain V r
4	1197	42.2	255	4 S31866	Ig gamma-1 chain C
5	1152	40.6	234	2 PT0207	Ig gamma chain C r
6	1148	40.4	326	1 G2HU	Ig gamma-2 chain C
7	1145	40.3	377	2 A23511	Ig gamma-3 chain C
8	1143	40.3	377	2 A60764	Ig gamma-3 chain C
9	1134.5	40.0	327	1 G4HU	Ig gamma-4 chain C
10	1116.5	39.3	289	1 G3HUI	Ig gamma-3 heavy c
11	921	32.4	323	1 GHRB	Ig gamma chain C r
12	913	32.2	328	2 I47160	Ig gamma 2b chain
13	913	32.2	328	2 I47159	Ig gamma 2a chain
14	906	31.9	277	2 I47162	Ig gamma 4 chain c
15	893.5	31.5	328	2 I47158	Ig gamma 1 chain c
16	884.5	31.2	329	1 G2GP	Ig gamma-2 chain C
17	881	31.0	328	2 I47161	Ig gamma-3 chain c
18	857.5	30.2	470	2 S22080	Ig heavy chain pre
19	849.5	29.9	308	2 C30554	Ig heavy chain C r
20	849.5	29.9	472	2 S31459	Ig gamma-1 chain -
21	841	29.6	329	1 G3MSC	Ig gamma-3 chain C
22	838.5	29.5	444	2 PC4436	monoclonal antibod
23	834.5	29.4	326	2 PS0017	Ig gamma-1 chain C
24	830	29.2	398	1 G3MSM	Ig gamma-3 chain C
25	826.5	29.1	324	1 G1MS	Ig gamma-1 chain C
26	826.5	29.1	333	2 PS0018	Ig gamma-2b chain
27	821.5	28.9	329	2 S00847	Ig gamma-2c chain
28	821.5	28.9	393	1 G1MSM	Ig gamma-1 chain C
29	817.5	28.8	322	2 PS0019	Ig gamma-2a chain

## ALIGNMENTS

### RESULT 1

NBHUIA

platelet glycoprotein Ib alpha chain precursor - human

N:Alternate names: membrane glycoprotein Ib alpha chain

N:Contains: glycosialin

C:Species: Homo sapiens (man)

C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 22-Jun-1999

C/Accession: A94174; A60435; A94173; S16945; I55355; A27075; A27102

R/Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hagen, F.S.; Papayannopoulou, T.; Roth, G.J

Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987

A>Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembran

A:Reference number: A94174; MUID:87289655; PMID:3303030

A:Accession: A94174

A:Molecule type: mRNA

A:Residues: 1-626 <LOP>

A:Cross-references: GB:J02940; NID:gi83499; PIDN:AAA52595.1; PID:G306793

R/Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemetson, I

Thromb. Haemost. 61, 448-453, 1989

A>Title: Isolation and characterization of human blood platelet mRNA and construction

d cloning of a GPIb coding cDNA insert.

A:Reference number: A60435; MUID:90020160; PMID:2799758

A:Accession: A60435

A:Molecule type: mRNA

A:Residues: 207-467 <WIC>

R/Titani, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.

Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987

A>Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet

A:Reference number: A94173; MUID:87289654; PMID:3497398

A:Accession: A94173

A:Molecule type: protein

A:Residues: 17-315 <TIT>

R/Hess, D.; Schaller, J.; Rickli, E.E.; Clemetson, K.J.

Eur. J. Biochem. 199, 389-393, 1991

A>Title: Identification of the disulphide bonds in human platelet glycosialin.

A:Reference number: S16945; MUID:91301149; PMID:2070794

A:Accession: S16945

A>Status: preliminary

A:Molecule type: protein

A:Residues: 224-227;262-270;277-282 <HES>

R/Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.

J. Biol. Chem. 267, 10055-10061, 1992

A>Title: Polymorphism of human glycoprotein Ib alpha results from a variable number o

ations.

A:Reference number: I55355; MUID:92250564; PMID:1577776

A:Accession: I55355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 412-427 <RES>

A:Cross-references: GB:S34436; NID:G249176; PIDN:AB22152.1; PID:G249177

A>Note: variant D

C:Comment: Glycoprotein Ib (GPIb), a surface membrane protein of platelets, participa

C:Comment: Platelet activation apparently involves disruption of the macromolecular C

C;Comment: Binding sites for von Willebrand factor and thrombin (the latter site with un  
C;Comment: Glycoalcenin, which is approximately coextensive with the extracellular part  
C;Genetics:  
A;Gene: GDB:GPIBA; GPIB  
A;Cross-references: GDB:118806; OMIM:231200  
A;Map position: 17pter-17p12  
C;Complex: heterodimer with platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein 2  
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein 2  
C;Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem repe  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MPT>  
F;48-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F;72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F;94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F;117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F;141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F;165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F;189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F;379-430/Region: proline/threonine-rich 9-residue repeats  
F;502-540/Domain: transmembrane #status predicted <TRM>  
F;541-626/Domain: intracellular #status predicted <INT>  
F;37,175/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;308/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match 56.9%; Score 1615.5; DB 1; Length 626;  
Best Local Similarity 84.2%; Pred. No. 1.2e-102;  
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy	1	MPLLLLLLPSPHHPHPICEVSKVASHLEVNCNKRNLTPALPDPKDTTLHLSENLLY	60
Db	1	MPLLLLLLPSPHHPHPICEVSKVASHLEVNCNKRNLTPALPDPKDTTLHLSENLLY	60
Qy	61	TFSLATMPYTRLQNLDRCELTKLOVDGTLPLVGLTLDLSHNOQLSLPGLQTLPALTY	120
Db	61	TFSLATMPYTRLQNLDRCELTKLOVDGTLPLVGLTLDLSHNOQLSLPGLQTLPALTY	120
Qy	121	LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL	180
Db	121	LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL	180
Qy	181	AGLLNGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPMLCNCEILYFRRWLQDNA	240
Db	181	AGLLNGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPMLCNCEILYFRRWLQDNA	240
Qy	241	ENVVVKQVVDVAVTNSVAVQCDNSDKPKVYKPKGCGPTLGDGDTLDLYYPEEDT	300
Db	241	ENVVVKQVVDVAVTNSVAVQCDNSDKPKVYKPKGCGPTLGDGDTLDLYYPEEDT	300
Qy	301	EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLPPK--	330
Db	301	EGDKVRATRTVVKFPTKAHTTTPMGLFYSWSTASLDQMPSSLHPTQSTKEQTTPPRWT	360
Qy	331	PKDTL-----MISTPEVT	344
Db	361	PNFTLHMESTFSPKST	379

RESULT 2  
GHU

Ig gamma-1 chain C region - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1981 #sequence revision 18-Aug-1982 #text change 16-Jul-1999  
C;Accession: A93433; S36861; S33887; B90563; A90564; A91723; A02146  
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.  
A;Reference number: A93433; MUID:82274238; PMID:6287432  
A;Accession: A93433  
A;Molecule type: DNA  
A;Residues: 1-330 <ELL>  
A;Cross-references: EMBL:Z17370  
A;Note: This sequence has the Gim(17) allotypic marker, 97-Lys, and the Gim(1) markers.  
A;Note: Lys-330 is removed after translation

R;Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A;Reference number: S33904  
A;Accession: S36861  
A;Molecule type: DNA  
A;Residues: 2-330 <HAR>  
A;Cross-references: EMBL:Z17370  
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of  
A;Reference number: S33887; MUID:83001943; PMID:6811139  
A;Accession: S33887  
A;Molecule type: DNA  
A;Residues: 88-113/235-330 <TAK>  
A;Cross-references: EMBL:Z17370  
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman  
Biochemistry 9, 3161-3170, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq  
A;Reference number: A90563; MUID:71064024; PMID:5489771  
A;Accession: B90563  
A;Molecule type: protein  
A;Residues: 1-96, 'R', '98-135 <CUN>  
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se  
A;Reference number: A90564; MUID:71064025; PMID:5530842  
A;Accession: Eu  
A;Accession: A90564  
A;Molecule type: protein  
A;Residues: 136-154, 'Q', '156-165, 'Q', '167-176, 'Q', '178-194, 'N', '196-197, 'D', '199-238, 'E', '2  
A;Note: this sequence has the Gim(non-1) markers, 239-Glu and 241-Met  
R;Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A;Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni  
igen Primärstruktur.  
A;Reference number: A91668; MUID:77070269; PMID:826475  
A;Accession: myeloma protein Nie  
A;Accession: B91668  
A;Molecule type: protein  
A;Residues: 1-34, 'Q', '36-96, 'K', '98-115, 'Q', '117-197, 'D', '199-238, 'D', '240, 'L', '242-268, 'E', 'E'  
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A;Title: Die Primärstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1  
A;Reference number: A91723; MUID:83289131; PMID:6884994  
A;Accession: myeloma protein KOL; disulfide bonds  
A;Accession: A91723  
A;Molecule type: protein  
A;Residues: 1-96, 'R', '98-197, 'D', '199-238, 'E', '240, 'M', '242-266, 'D', '268-271, 'D', '273-330 <  
A;Note: this sequence has the Gim(3) and Gim(non-1) markers  
R;Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disul  
A;Reference number: A90565; MUID:71064027; PMID:4923144  
A;Accession: annotation; disulfide bonds  
R;Brekke, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunog  
enamide cleavage products, and the disulfide bridges.  
A;Reference number: A91667; MUID:77070267; PMID:1002129  
A;Accession: annotation; disulfide bonds  
C;Genetics:  
A;Gene: GDB:IGHG1  
A;Cross-references: GDB:120085; OMIM:147100  
A;Map position: 14q32.33-14q32.33  
A;Introns: 9/1; 114/1; 224/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L)  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;20-85/Domain: immunoglobulin homology <IM1>

```

QY      426 QPREQVYTLPSREEMTKNQVSLTCLVKGYFSPDIAVEMESNGOPENNYKTTTPVLDS    485
DB      269 QPREQVYTLPSREEMTKNQVSLTCLVKGYFSPDIAVEMESNGOPENNYKTTTPVLDS    328
QY      486 GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK    531
DB      329 GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK    374

RESULT 4
S31866
Ig gamma-1 chain C region - synthetic
C:Species: synthetic
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
A:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
A:Accession: S31866
R:Pilpula, D.
submitted to the EMBL Data Library, February 1993
A:Description: Screening method for protein-protein interactions of cloned gene products
A:Reference number: S31866
A:Accession: S31866
A:Molecule type: mRNA
A:Residues: 1-255 <PIL>
A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match          42.2%; Score 1197; DB 4; Length 255;
Best Local Similarity 96.9%; Pred.No. 1.4e-74;
Matches 219; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      306 RPTCPPCPAPEALGAPSVLPFPKPCKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYVDG    365
DB      30 KWTCTCCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYVDG    89
QY      366 VEVHNAKTVPREQNSTYRVVSVLTVLHQDWLNKEYKCKVSNKALPAPIEKTISKAKG    425
DB      90 VEVHNAKTVPREQNSTYRVVSVLTVLHQDWLNKEYKCKVSNKALPAPIEKTISKAKG    149
QY      426 QPREQVYTLPSREEMTKNQVSLTCLVKGYFSPDIAVEMESNGOPENNYKTTTPVLDS    485
DB      150 QPREQVYTLPSREDELTKNQVSLTCLVKGYFSPDIAVEMESNGOPENNYKTTTPVLDS    209
QY      486 GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK    531
DB      210 GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK    255

RESULT 5
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C:Accession: PT0207
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A:title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MOID:91287716; PMID:2062315
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <EHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match          40.6%; Score 1152; DB 2; Length 234;
Best Local Similarity 93.8%; Pred.No. 1.5e-71;
Matches 213; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY      304 KYVRP-----HTCPRCPAPEALGAPSVLPFPKPCKDTLMISRTPEVTCVVVDVSHEDPEV    357
DB      8 KYRKSCDTHTCRCACAFELTGGSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEV    67

```



QY 358 KENNYVGVGVHNAKTPREOVNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPVPIE 417  
 Db 68 KFNYYVGVGVHNAKTPREOVNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPVPIE 127  
 QY 418 KTIKAKGQPREQVYTLTPSRDELTQVSLTCLVKGFYPSDIAVEMESNGOPENNYKT 477  
 Db 128 KTIKAKGQPREQVYTLTPSRDELTQVSLTCLVKGFYPSDIAVEMESNGOPENNYKT 187  
 QY 478 TPVLDSDGFFLYSLKTVDKSRWQGNVPSCSVMHEALHNHYTQKS 524  
 Db 188 TPVLDSDGFFLYSLKTVDKSRWQGNVPSCSVMHEALHNHYTQKS 234

RESULT 6  
 G2HU  
 Ig gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con-  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA  
 A:Residues: 1-326 <ELL>  
 A:Cross-references: GB:V00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g6066056  
 A:Note: Lys-326 is probably removed posttranslationally  
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
 J. Immunol. 125, 1048-1054, 1980  
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and e-  
 A:Reference number: A92809; MUID:81007873; PMID:6774012  
 A:Contents: myeloma protein T11  
 A:Accession: A92809  
 A:Molecule type: protein  
 A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
 A:Note: Trp-156 is at or near the complement-binding site  
 R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
 Can. J. Biochem. 57, 758-767, 1979  
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
 A:Reference number: A90752; MUID:80001357; PMID:113060  
 A:Contents: myeloma protein Zie  
 A:Accession: A90752  
 A:Molecule type: protein  
 A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
 A:Note: this sequence has since been revised  
 R:Hofmann, T.; Parr, D.M.  
 Mol. Immunol. 16, 923-925, 1979  
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
 A:Reference number: A93132; MUID:80114419; PMID:118920  
 A:Contents: Zie  
 A:Accession: A93132  
 A:Molecule type: protein  
 A:Residues: 238-275 <HOF>  
 R:Hofmann, T.; Parr, D.M.  
 submitted to the Atlas, March 1980  
 A:Reference number: A94591  
 A:Contents: annotation; Zie  
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
 ned  
 R:Milstein, C.; Frangione, B.  
 Biochem. J. 121, 217-225, 1971  
 A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
 A:Reference number: A90253; MUID:72033500; PMID:4940472  
 A:Contents: annotation; myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.  
 A:Reference number: A93157; MUID:69064124; PMID:5782707  
 A:Contents: annotation; Sa, disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (co light chain) #status experimental  
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (co heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1148; DB 1; Length 326;  
 Best Local Similarity 72.2%; Pred. No. 4.4e-71;  
 Matches 226; Conservative 18; Mismatches 33; Indels 36; Gaps 6;

QY 253 KAVTSNVAHVQDSDSKFP-----VYKYP-----GKGCPYL----- 283  
 Db 16 RSTSESTAALGCLVKDYFPEPTVYSNMSGALTSGVHTFPAVLQSSGLYSLSSVVPSSN 75  
 QY 284 -GDRGDTLDYIPEEDTEGDKVRPH-----TCPCPAPEALGAPSVLFPPKPKDTLMIS 338  
 Db 76 FGTQYTCNVDHKP-SNTKVDKTVKRCVCCPCPPAVAG-PSVLFPPKPKDTLMIS 133  
 QY 339 RPEVTCVVDVSHEDPEVKENWYGVGVENAKTKPREOYNSTYRVVSVLTVLHODWL 398  
 Db 134 RPEVTCVVDVSHEDPEVQFNWYGVGVENAKTKPREOYNSTYRVVSVLTVLHODWL 193  
 QY 399 NGKEYKCKVSNKALPVPIEKTISKAKGQPREPQVYTLTPSRDELTQVSLTCLVKGFY 458  
 Db 194 NGKEYKCKVSNKALPVPIEKTISKAKGQPREPQVYTLTPSRDELTQVSLTCLVKGFY 253  
 QY 459 SDIAVEMESNGOPENNYKTTPVLDSDGFFLYSLKTVDKSRWQGNVPSCSVMHEALHN 518  
 Db 254 SDIAVEMESNGOPENNYKTTPVLDSDGFFLYSLKTVDKSRWQGNVPSCSVMHEALHN 313  
 QY 519 HYTKSLSLSPCK 531  
 Db 314 HYTKSLSLSPCK 326

RESULT 7  
 A23511  
 Ig gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 40.3%; Score 1145; DB 2; Length 377;  
 Best Local Similarity 64.4%; Pred. No. 8.7e-71;  
 Matches 230; Conservative 17; Mismatches 56; Indels 54; Gaps 7;

QY 194 LIQENSLY-----TIPKGFQSHLLPFAFLHGNPWLNCCELLYPRRWLQNAENYVWK 247  
 Db 56 VLQSSGLYSLSSVWTVPSSSLGT-----QTYTCN-----VNHKPSNTKVDK 96



A:Reference number: A90442; MUID:81021548; PMID:6774747  
A:Contents: heavy chain disease protein Wis  
A:Accession: A90442  
A:Molecule type: protein  
A:Residues: 1-289 <PRA>  
A>Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain  
A>Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 cd  
A>Note: the sequence of residues 42-76 was taken from the reference that follows  
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
A:Reference number: A92219; MUID:77118561; PMID:402363  
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W  
A:Accession: A92219  
A:Molecule type: protein  
A:Residues: 12-97 <MIC>  
A>Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
A>Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
A>Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter  
R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A:Title: The amino acid sequence of 'heavy chain disease' protein ZUC. Structure of the  
A:Reference number: A90198; MUID:77021516; PMID:823945  
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues  
A:Accession: A90198  
A:Molecule type: protein  
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A>Note: this protein lacks most of the V region, all of the CH1 region, and part of the  
R:Alexander, A.; Steinmetz, M.; Barricault, D.; Frangione, B.; Franklin, E.C.; Hood, L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion  
A:Reference number: A93915; MUID:82247835; PMID:6808505  
A:Contents: heavy chain disease protein Omm  
A:Accession: A93915  
A:Molecule type: mRNA  
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157  
A>Note: a carboxyl-terminal Lys is removed posttranslationally  
A>Note: this sequence may represent an allelic form or another gamma chain subclass  
C:Comment: The heavy chain disease protein Wis is shown.  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid  
F:203-270/Domain: immunoglobulin homology <IM>  
F:1/Modified site: pyroglutamic acid (Gln) #status experimental  
F:16,140/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 39.3%; Score 1116.5; DB 1; Length 289;  
Best Local Similarity 78.7%; Pred. No. 5.3e-69;  
Matches 210; Conservative 11; Mismatches 21; Indels 25; Gaps 2;  
QY 264 CDNSDKFVYKPGKCGPTLGDGDTLDYYPEDTEGDKVRPHTCPPCPAPEALGAPS 323  
Db 48 CDTPPCPCPEP-KSCDT-----PPPCPCPEALGAPS 82  
QY 324 VLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKPNVYDGVFVHNAKTPREEVNST 363  
Db 83 VLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKPNVYDGVFVHNAKTPREEVNST 142  
QY 384 YRVSVLTVLHODWLNKGKCKYKCKVSNKALPVPPIEKTISKAGQPREPQVYTLPPSREMT 443  
Db 143 FRVSVLTVLHONWLDGKGYKCKYKCKVSNKALPVPPIEKTISKAGQPREPQVYTLPPSREMT 202  
QY 444 KNOVSLTCLVNGFVPSDIATVFNWESNGQPNKYKTPPVLSDSGSFFLYSKLTVDKSRWQ 503  
Db 203 KNOVSLTCLVNGFVPSDIATVFNWESNGQPNKYKTPPVLSDSGSFFLYSKLTVDKSRWQ 262  
QY 504 GNVSFCSVWHEALHNHYTKSLSPG 530  
Db 263 GNVSFCSVWHEALHNHYTKSLSPG 289

RESULT 11  
GHRB  
Ig gamma chain C region - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A:Title: Nucleotide sequence of a rabbit IGH heavy chain from the recombinant F-I hap  
A:Reference number: A91749; MUID:84030930; PMID:6313520  
A:Accession: A91749  
A:Molecule type: mRNA  
A:Residues: 1-323 <BER>  
A>Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-  
R:Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog  
A:Reference number: A90290; MUID:76135469; PMID:1243651  
A:Accession: A90290  
A:Molecule type: protein  
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A:Title: Heavy chain genes of rabbit IGH; isolation of a cDNA encoding gamma heavy ch  
A:Reference number: A93928; MUID:83299917; PMID:6193512  
A:Accession: A93928  
A:Molecule type: mRNA  
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112  
A>Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma  
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul  
A:Reference number: A90245; MUID:70110015; PMID:5461106  
A:Accession: A90245  
A:Molecule type: protein  
A:Residues: 132-143, 'E', 145-161 <PRU>  
R:Hall, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse  
A:Reference number: A94416  
A:Accession: A94416  
A:Molecule type: protein  
A:Residues: 129-131, 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '  
A>Note: this has the e15 allotypic marker, 185-Ala  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( chain  
disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-82/Domain: immunoglobulin homology <IM1>  
F:130-199/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 32.4%; Score 921; DB 1; Length 323;  
Best Local Similarity 61.1%; Pred. No. 1.4e-55;  
Matches 173; Conservative 35; Mismatches 45; Indels 30; Gaps 3;  
QY 252 VKATSVASVQCDNSDKFVYKYPYKPGKCGPTLGDGDTLDYYPEDTEGDK-VRPHTC 310  
Db 68 VVSVTSQPPVTCNA-----HPATNTKVDKTVAPSTC 100  
QY 311 --PPCPAPEALGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEV 368  
Db 101 SKTPCPPELLGGSVIFPPKPKDLMISRTPEVTCVVVDVSDDEVOFTWYINNEQV 160  
QY 369 HNATKTPREQYNSTYRVSVLTVLHODWLNKGKCKYKCKVSNKALPVPPIEKTISKAGQPR 428  
Db 161 RTARPLREQQFNSTIRVSVLTPLTHQDWLRGKGFCKVHNAKLPAPIEKTISKARQPL 220  
QY 429 EPOVYTLPPSREMTKNQVSLTCLVKGFPVSDIATVFNWESNGQPNKYKTPPVLSDSGSF 498  
Db 221 EPKYVTMGPPREELSSRSVSLTCMNGFYPSDISVSEWEKNGKAEDNYKTTTPAVLSDSGSY 280



C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match	31.5%	Score	893.5	DB 2	Length	328			
Best Local Similarity	68.7%	Pred. No. 1.le-53							
Matches	167	Conservative	29	Mismatches	40	Indels	7	Gaps	3

  

QY	295	YPEDTEGDK----	VRPHTCPGCPAPEALCAPSVFLFPKPKDTLMISRTPEVTCVWVDV	350
DB	87	HPATTTKVDKRVGIHQPTCPCPGCEVAG-PSVFIFPPKPKDTLMISQTPPEVTCVWVDV	145	
QY	351	SHEDPEVKEVWYVDGVEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNK	410	
DB	146	SKEHAEVQFSWYVDGVEVHTAETRPKEEQFNSTYRVSVLPQHODWLKGEKCKVNV	205	
QY	411	ALPVPTEKTISKAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFSPSDIAVEWESNGQ	470	
DB	206	DLPAITRTISKAGQPREPOVYTLPPPAEELSRSKVTLTCLVIGFYPPDIHVEWKSNGQ	265	
QY	471	--PENNYKTTTPVLDSGSFELYSKLTVDKSRWQQGVFSCSYMHEALHNHYTOKSLSL	528	
DB	266	PEPENTYRTTTPPOODVDGTFFLYSKLAVDKARWDHGDGKFECAVMHEALHNHYTOKSISKT	325	
QY	529	PGK	531	
DB	326	QK	328	

Search completed: October 14, 2003, 06:17:18  
Job time : 25.6977 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 14.8186 Seconds  
(without alignments)  
1685.124 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPLLLLLLLLPSLPHPIC.....MHEALHNHYTKSLSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615.5	56.9	626	1 GPBA_HUMAN	P07359 homo sapien
2	1205	42.4	330	1 GC1_HUMAN	P01857 homo sapien
3	1148	40.4	326	1 GC2_HUMAN	P01859 homo sapien
4	1134.5	40.0	327	1 GC4_HUMAN	P01861 homo sapien
5	1121.5	39.5	290	1 GC3_HUMAN	P01860 homo sapien
6	921	32.4	323	1 GC_RABIT	P01870 oryctolagus
7	894.5	31.2	329	1 GC2_CAVPO	P01862 cavia porce
8	841	29.6	329	1 GC3_MOUSE	P22436 mus musculu
9	834.5	29.4	326	1 GC1_RAT	P20759 rattus norv
10	830	29.2	398	1 GC3M_MOUSE	P03987 mus musculu
11	826.5	29.1	324	1 GC1_MOUSE	P01868 mus musculu
12	826.5	29.1	333	1 GCB_RAT	P20761 rattus norv
13	821.5	28.9	329	1 GCC_RAT	P20762 rattus norv
14	821.5	28.9	393	1 GC1M_MOUSE	P01869 mus musculu
15	817.5	28.8	322	1 GCA_RAT	P20760 rattus norv
16	809	28.3	330	1 GCAB_MOUSE	P01863 mus musculu
17	804.5	28.3	335	1 GCAB_MOUSE	P01864 mus musculu
18	804	28.3	399	1 GCAM_MOUSE	P01865 mus musculu
19	770	27.1	336	1 GCB_MOUSE	P01866 mus musculu
20	765	26.9	405	1 GCB_MOUSE	P01867 mus musculu
21	363	12.8	421	1 EPC_MOUSE	P06336 mus musculu
22	360	12.7	391	1 MUCB_HUMAN	P04220 homo sapien
23	359.5	12.7	429	1 EPC_RAT	P01855 rattus norv
24	359	12.6	454	1 MUC_HUMAN	P01871 homo sapien
25	352	12.4	455	1 MUC_MOUSE	P01872 mus musculu
26	342	12.0	476	1 MUCM_MOUSE	P01873 mus musculu
27	341	12.0	428	1 EPC_HUMAN	P01854 homo sapien
28	338	11.9	458	1 MUC_RABIT	P03988 oryctolagus
29	331.5	11.7	450	1 MUC_CANFA	P01874 canis famil
30	329	11.6	454	1 MUC_MESAU	P06337 mesocricetu
31	328	11.6	479	1 MUCM_RABIT	P04221 oryctolagus
32	322	11.3	457	1 MUC_SUNMU	P20768 suncus muri
33	301	10.6	347	1 AZGL_HUMAN	P02750 homo sapien

#### RESULT 1

ID	GPBA_HUMAN	STANDARD;	PRT;	626 AA.
AC	P07359; Q9HDC7;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DE	Platelet glycoprotein Ib alpha chain precursor (Glycoprotein Ibalpha)			
DE	(GP-Ib alpha) (GP1BA) (GP1b-alpha) (CD42B-alpha) (CD42B) [Contains:			
DE	Glycocalicin].			
GN	GP1BA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI TaxID=9606;			
RN	[1] _SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=87289655; PubMed=3303030;			
RA	Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T.,			
RA	Roth G.J.;			
RT	"Cloning of the alpha chain of human platelet glycoprotein Ib: a			
RT	transmembrane protein with homology to leucine-rich alpha 2-			
RT	glycoprotein";			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).			
RP	SEQUENCE FROM N.A., AND VARIANT PHE-86.			
RX	MEDLINE=22034223; PubMed=12038791;			
RA	Matsubara Y., Murata M., Moriki T., Yokoyama K., Watanabe N.,			
RA	Nakajima H., Hada M., Kawano K., Aoki N., Yoshino H., Ikeda Y.;			
RT	"A novel polymorphism, 70Leu/Phe, disrupts a consensus Leu residue			
RT	within the leucine-rich repeat sequence of platelet glycoprotein			
RT	Ibalpha.";			
RL	Thromb. Haemost. 87:867-872(2002).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT MET-161.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,			
RA	Nickerson D.A.;			
RN	[5]			
RP	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE OF 17-315.			
RX	MEDLINE=87289654; PubMed=3497398;			
RA	Titani K., Takio K., Handa M., Ruggeri Z.M.;			
RT	"Amino acid sequence of the von Willebrand factor-binding domain of			
RT	platelet membrane glycoprotein Ib.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).			
RP	DISULFIDE BONDS.			
RX	MEDLINE=91301149; PubMed=2070794;			
RA	Hess D., Schaller J., Rickli E.E., Clemetson K.J.;			

34	296	10.4	438	1	HVC2_HETFR	P23085 heterodontu
35	289.5	10.2	446	1	MUC_CHICK	P01875 gallus gall
36	286	10.1	438	1	HVC5_HETFR	P23087 heterodontu
37	284.5	10.0	299	1	ALC_RABIT	P01879 oryctolagus
38	282.5	10.0	353	1	ALC1_HUMAN	P01876 homo sapien
39	277	9.8	461	1	HVCM_HETFR	P23088 heterodontu
40	275	9.7	393	1	HVC3_HETFR	P23086 heterodontu
41	274	9.7	340	1	ALC2_HUMAN	P01877 homo sapien
42	274	9.7	370	1	HVC1_HETFR	P23084 heterodontu
43	273.5	9.6	353	1	ALC1_GORGO	P20758 gorilla gor
44	266	9.4	440	1	OMGP_MOUSE	Q63912 mus musculu
45	265.5	9.4	646	1	FLR1_HUMAN	Q9nzul homo sapien

#### ALIGNMENTS

RT "Identification of the disulphide bonds in human platelet  
 RL glycofibrinogen";  
 RN Sur. J. Biochem. 199:389-393(1991).  
 RX [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
 RX MEDLINE=22171415; PubMed=12183630;  
 RA Huizinga E.G., Tsuji S., Romijn R.A., Schiphorst M.E., de Groot P.G.,  
 RA Sixma J.J., Gros P.;  
 RT "Structures of glycoprotein Ib alpha and its complex with von  
 RT Willebrand factor A1 domain";  
 RL Science 297:1176-1179(2002).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 17-304, AND SULFATION OF  
 RP TYR-292; TYR-294 AND TYR-295.  
 RX MEDLINE=22217968; PubMed=12087105;  
 RA Uff S., Clemetson J.M., Harrison T., Clemetson K.J., Emsley J.;  
 RT "Crystal structure of the platelet glycoprotein Ib alpha N-terminal  
 RT domain reveals an unmasking mechanism for receptor activation";  
 RL J. Biol. Chem. 277:35657-35663(2002).  
 RN [9]  
 RP VARIANT SIBA MET-161.  
 RX MEDLINE=92265982; PubMed=1586750;  
 RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,  
 RA Ruggeri Z.M.;  
 RT "Genetic and structural characterization of an amino acid dimorphism  
 RT in glycoprotein Ib alpha involved in platelet transfusion  
 RT refractoriness";  
 RL Blood 79:3086-3093(1992).  
 RN [10]  
 RP VARIANT BSS PHE-73.  
 RX MEDLINE=92110577; PubMed=1730088;  
 RA Miller J.L., Lyle V.A., Cunningham D.;  
 RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein  
 RT Ib alpha leucine tandem repeat occurring in patients with an  
 RT autosomal dominant variant of Bernard-Soulier disease";  
 RL Blood 79:439-446(1992).  
 RN [11]  
 RP VARIANT BSS VAL-172.  
 RX MEDLINE=93388951; PubMed=7690774;  
 RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,  
 RA De Marco L., Ruggeri Z.M.;  
 RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib  
 RT alpha resulting in the Bernard-Soulier syndrome";  
 RL J. Clin. Invest. 92:1213-1220(1993).  
 RN [12]  
 RP VARIANT BSS SER-225.  
 RX MEDLINE=95118882; PubMed=7819107;  
 RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,  
 RA Ribera A., Gallardo D.;  
 RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha  
 RT gene is associated with Bernard-Soulier syndrome";  
 RL Br. J. Haematol. 88:839-844(1994).  
 RN [13]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RX MEDLINE=91271273; PubMed=2052556;  
 RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
 RT "Mutation in the gene encoding the alpha chain of platelet  
 RT glycoprotein Ib in platelet-type von Willebrand disease";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765(1991).  
 RN [14]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RX MEDLINE=93253059; PubMed=8486780;  
 RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
 RT "Expression of the phenotypic abnormality of platelet-type von  
 RT Willebrand disease in a recombinant glycoprotein Ib alpha fragment";  
 RL J. Clin. Invest. 91:2133-2137(1993).  
 RN [15]  
 RP VARIANT PSEUDO-VWD VAL-255.  
 RX MEDLINE=93214031; PubMed=8384898;  
 RA Russell S.D., Roth G.J.;  
 RT "Pseudo-von Willebrand disease: a mutation in the platelet  
 RT glycoprotein Ib alpha gene associated with a hyperactive surface  
 RT receptor";

Blood 81:1787-1791(1993).  
 [16]  
 RL VARIANT BSS LEU-195 DEL.  
 RX MEDLINE=95178321; PubMed=7873390;  
 RA de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
 RA Chevalier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
 RT "A three-base deletion removing a leucine residue in a leucine-rich  
 RT repeat of platelet glycoprotein Ib alpha associated with a variant of  
 RT Bernard-Soulier syndrome (Nancy 1)";  
 RL Br. J. Haematol. 89:386-396(1995).  
 RN [17]  
 RP VARIANT BSS ARG-81.  
 RX MEDLINE=98303759; PubMed=9639514;  
 RA Kenny D., Jonsson O.G., Morawetz P.A., Montgomery R.R.;  
 RT "Naturally occurring mutations in glycoprotein Ib alpha that result in  
 RT defective ligand binding and synthesis of a truncated protein";  
 RL Blood 92:175-183(1998).  
 RN [18]  
 RP VARIANTS HIS-72 AND MET-161.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [19]  
 RP ERRATUM.  
 RX Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [20]  
 RP VARIANT BSS PRO-145.  
 RX MEDLINE=99189763; PubMed=10089893;  
 RA Koskela S., Partanen J., Salmi T.T., Kekomaki R.;  
 RT "Molecular characterization of two mutations in platelet glycoprotein  
 RT (GP) Ib alpha in two Finnish Bernard-Soulier syndrome families";  
 RL Eur. J. Haematol. 62:160-168(1999).  
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,  
 CC participates in the formation of platelet plugs by binding to the  
 CC A1 domain of von Willebrand factor, which is already bound to the  
 CC subendothelium.  
 CC -!- SUBUNIT: Heterodimer composed of GP-Ib alpha and beta; disulfide  
 CC linked. GP-IX is complexed with the GP-Ib heterodimer via a non  
 CC covalent linkage.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- PTM: Glycosylation, which is approximately coextensive with the  
 CC extracellular part of the molecule, is cleaved off by calpain  
 CC during platelet lysis.  
 CC -!- POLYMORPHISM: Position 161 is associated with platelet-specific  
 CC alloantigen Siba. Siba(-) has Thr-161 and Siba(+) has Met-161.  
 CC Siba is involved in neonatal alloimmune thrombocytopenia (NATP).  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of Bernard-Soulier  
 CC syndrome (BSS). BSS patients have unusually large platelets and  
 CC have a clinical bleeding tendency.  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of von Willebrand  
 CC disease (vWD) known as platelet-type von Willebrand disease or  
 CC pseudo-von Willebrand disease. This autosomal dominant bleeding  
 CC disorder is caused by an increased affinity of GP-Ib for soluble  
 CC vWF resulting in impaired hemostatic function due to the removal  
 CC of vWF from the circulation.  
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption  
 CC of the macromolecular complex of GP-Ib with the platelet  
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-  
 CC binding protein.  
 CC -!- MISCELLANEOUS: Binding sites for von Willebrand factor and  
 CC thrombin (the latter site with unknown function) are in the amino-  
 CC terminal part of the molecule.  
 CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.

Query Match 56.9%; Score 1615.5; DB 1; Length 626;  
 Best Local Similarity 84.2%; Pred. No. 8.8e-108;  
 Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

QY 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCNLTALPPDLPKOTITLHLSNLLY 60  
 DB 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCNLTALPPDLPKOTITLHLSNLLY 60

QY 61 TFSLATLMPYTRLTQNLNLCORCELTQVGTGLPVLTGLDLSHNQSLPLGQOTIPALTV 120  
 DB 61 TFSLATLMPYTRLTQNLNLCORCELTQVGTGLPVLTGLDLSHNQSLPLGQOTIPALTV 120

QY 121 LDVSFNRLTSLPGALRGELGELQELKGNELKTLTPPGGLTTPPKLEKLSLNNLTLP 180  
 DB 121 LDVSFNRLTSLPGALRGELGELQELKGNELKTLTPPGGLTTPPKLEKLSLNNLTLP 180

QY 181 AGLLNGLENDTLTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRRMLQDNA 240  
 DB 181 AGLLNGLENDTLTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRRMLQDNA 240

QY 241 ENYVYWKQVYKAVTSNVASVOCNSDKFPVYKYPKSCPTLGDGDTLDLYDYPEEDT 300  
 DB 241 ENYVYWKQVYKAVTSNVASVOCNSDKFPVYKYPKSCPTLGDGDTLDLYDYPEEDT 300

QY 301 EGDQVRA-----PHTCP-----PCPA---PEALGAPSVFLPPPK-- 330  
 DB 301 EGDQVRA-----PHTCP-----PCPA---PEALGAPSVFLPPPK-- 330

QY 331 PKDTL-----MISRTPEVT 344  
 DB 361 PNFTLHMSITFSKTPKST 379

## RESULT 2

GC1\_HUMAN STANDARD; PRT; 330 AA.

AC P01857;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-1 chain C region.  
 GN IGHG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82274238; PubMed=6287432;  
 RA Ellison J.W., Berson B.J., Hood L.E.;  
 RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";  
 RL Nucleic Acids Res. 10:4071-4079(1982).  
 RN [2]  
 RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).  
 RX MEDLINE=71064024; PubMed=5489771;  
 RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,  
 RA Maxdal M.J., Edelman G.M.;  
 RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino  
 acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";  
 RL Biochemistry 9:3161-3170(1970).  
 RN [3]  
 RP SEQUENCE OF 136-329 (EU).  
 RX MEDLINE=71064025; PubMed=5530842;  
 RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,  
 RA Edelman G.M.;  
 RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino  
 acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";  
 RL Biochemistry 9:3171-3181(1970).  
 RN [4]  
 RP SEQUENCE (MYELOMA PROTEIN NIE).  
 RX MEDLINE=77070269; PubMed=826475;  
 RA Ponstingl H., Hilschmann N.;

RT "The rule of antibody structure. The primary structure of a  
 monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The  
 chymotryptic peptides of the H-chain, alignment of the tryptic  
 peptides and discussion of the complete structure.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).  
 RN [5]  
 RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.  
 RX MEDLINE=83289131; PubMed=6884994;  
 RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;  
 RT "Three-dimensional structure determination of antibodies. Primary  
 structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71064027; PubMed=4923144;  
 RA Gall W.E., Edelman G.M.;  
 RT "The covalent structure of a human gamma G-immunoglobulin. X.  
 Intrachain disulfide bonds.";  
 RL Biochemistry 9:3188-3196(1970).  
 RN [7]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 IgG1 immunoglobulin (myeloma protein Nie). I: Purification and  
 characterization of the protein, the L- and H-chains, the  
 cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 and its complex with fragment B of protein A from Staphylococcus  
 aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
 MARKER & THE GIM (NON-1) MARKERS.  
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 35,116,198,269 & 272.  
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 268-272.  
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 RESIDUES 198, 267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A93433; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR PDB; 1AJ7; 12-NOV-97.  
 DR PDB; 1DSB; 09-FEB-00.  
 DR PDB; 1DS1; 09-FEB-00.  
 DR PDB; 1D6V; 04-OCT-00.  
 DR PDB; 1DN2; 17-MAY-00.  
 DR PDB; 1S4K; 06-JUN-01.  
 DR PDB; 1FCC; 20-JUL-95.  
 DR PDB; 1H2H; 12-JUN-02.  
 DR PDB; 1H7Z; 08-AUG-01.  
 DR PDB; 1H1S; 16-MAY-01.  
 DR PDB; 1H1X; 16-MAY-01.  
 DR PDB; 1L6X; 10-APR-02.  
 DR PDB; 2RCS; 12-NOV-97.  
 DR Genew; HGNC:5525; IGHG1.





[16] SEQUENCE OF 238-275 (ZIE).  
 RX MEDLINE=80114419; PubMed=118920;  
 RA Hofmann T., Parr D.M.;  
 RT "A note of the amino acid sequence of residues 381-391 of human  
 immunoglobulin gamma chains.";  
 RL Mol. Immunol. 16:923-925(1979).  
 [17]  
 RN REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
 RA Hofmann T., Parr D.M.;  
 RL Submitted (MAR-1980) to the PIR data bank.  
 [18]  
 RN SEQUENCE OF 1-121 (DOT).  
 RX MEDLINE=95255298; PubMed=77371190;  
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
 RT "Characterization of the two unique human anti-flavin monoclonal  
 immunoglobulins.";  
 RL Eur. J. Biochem. 228:886-893(1995).  
 [19]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=72033500; PubMed=4940472;  
 RA Milstein C., Frangione B.;  
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
 RL Biochem. J. 121:217-225(1971).  
 [10]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=69064124; PubMed=5782707;  
 RA Frangione B., Milstein C., Pink J.R.L.;  
 RT "Structural studies of immunoglobulin G.";  
 RL Nature 221:145-148(1969).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J00230; AAB59393.1; ..  
 DR PIR; A93906; G2HU.  
 DR HSSP; P01857; IFCL.  
 DR Genew; HGNC:5526; IGHG2.  
 DR MIM; 147110; ..  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:003823; F:antigen binding activity; TAS.  
 DR GO; GO:0004955; P:immune response; NAS.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003597; IG-cl.  
 DR Pfam; PF00047; IG\_3.  
 DR SMART; SM00407; IGcl; 2.  
 DR PROSITE; PS00835; IG-LIKE; 3.  
 DR PROSITE; PS00390; IG\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT SITE 156 156 AT OR NEAR THE COMPLEMENT-BINDING SITE.  
 FT MOD\_RES 326 326 REMOVED POST-TRANSLATIONALLY (PROBABLE).  
 FT VARIANT 60 60 S -> A (IN MYELOMA PROTEINS TIL & ZIE).  
 FT CONFLICT 109 109 /FTID=VAR\_003889.  
 C -> S (IN REF. 3).

SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
 Query Match 40.4%; Score 1148; DB 1; Length 326;  
 Best Local Similarity 72.2%; Pred. No. 8.6e-75;  
 Matches 226; Conservative 18; Mismatches 33; Indels 36; Gaps 6;  
 Qy 253 KAVTSNVAQCDNSDKFP-----VYKYP-----GKGCPYL----- 283  
 Db 16 RSTSESTAALGCLVKDYFPEPTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSN 75  
 Qy 284 -GDEGDTLDYDYPEDTSGDKVRPH-----TPPCAPALGAPSVFLPPPKKDTLMIS 338  
 Db 76 FGTQTYTCNVDHKP-SNTKVDKTVKCKVCPCPPCPAPVAG-PSVFLFPPPKKDTLMIS 133  
 Qy 339 RTEVTCTVVVDVSHEDPEVKFNWYDGVGVHNAKTKPREOVNSTYRVSVLTVLHQQDL 398  
 Db 134 RTEVTCTVVVDVSHEDPEVQFNWYDGVGVHNAKTKPREEQFNSTFRVSVLTVLHQQDL 193  
 Qy 399 NGKEYKCKVSNKALPVIETKISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYP 458  
 Db 194 NGKEYKCKVSNKGLPAPIETKISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYP 253  
 Qy 459 SDIAVWESNGQPNENYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNVPSCSVMHEALHN 518  
 Db 254 SDIAVWESNGQPNENYKTPPMLDSDGSGFFLYSKLTVDKSRWQQGNVPSCSVMHEALHN 313  
 Qy 519 HYTKQKSLSLSPGK 531  
 Db 314 HYTKQKSLSLSPGK 326  
 RESULT 4  
 GC4\_HUMAN  
 ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
 AC P01861;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-4 chain C region.  
 GN IGHG4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8157104; PubMed=6299662;  
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";  
 RL DNA 1:11-18(1981).  
 RN [2]  
 RP SEQUENCE OF 1-30 AND 81-326.  
 RX MEDLINE=70207560; PubMed=4192699;  
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;  
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
 constant region of a gamma 4 chain.";  
 RL Biochem. J. 117:33-47(1970).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; K01316; AAB59394.1; ALT\_INIT.  
 DR PIR; A90933; G4HU.  
 DR PDB; 1ADO; 16-SEP-98.  
 DR Genew; HGNC:5528; IGHG4.  
 DR MIM; 147130; ..  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0003823; F:antigen binding activity; TAS.

```

DR GO: 0006955; P: immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS00835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 40.0%; Score 1134.5; DB 1; Length 327;
Best Local Similarity 71.2%; Pred. No. 7.9e-74;
Matches 222; Conservative 17; Mismatches 40; Indels 33; Gaps 3;

QY 253 KAVTSNVASVOCDSNPKP-----VYKYPG-----KGCP 282
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
16 RSTSESTAALGCLVKDYFPEPTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTPSS 75
QY LGDEGDTLDYDYPEEDTEGDKVRPH---TCPPCPAPEALGAPSVFLFPPKPKDTLMISR 339
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
76 LGTKTYTCNVDRHKPSNTKVDKRVESKYGPPCPAPFLGPGSVFLFPPKPKDTLMISR 135
QY TPVTCVVDVSHEDSEVKFNWTVGVVHNATKPKREQVSTYRVSVLTVLHQDWL 399
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
136 TPVTCVVDVSDQEDPEVQFNWTVGVVHNATKPKREQVSTYRVSVLTVLHQDWLN 195
QY 400 GREYKCKVSNKALPVPTEKTSKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGF 459
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
196 GREYKCKVSNKNGLPSSIEKTSKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGF 255
QY DIAVESNQCQPNNTKPTPVLVDSGSGFLYSLKLTVDKSRMQGQNVFSCVNGHEALHH 519
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
256 DIAVESNQCQPNNTKPTPVLVDSGSGFLYSLKLTVDKSRMQGQNVFSCVNGHEAL 315
QY 520 YTKSLSLSGK 531
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
316 YTKSLSLSGK 327

RESULT 5
GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
GNIGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RX MEDLINE=61021548; PubMed=6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308(1980).
RN [2]
RP REVISIONS, TO 12-97 (PROTEIN WIS).
RX MEDLINE=77118561; PubMed=402363;
Michaelson T.E., Frangione B., Franklin E.C.;
"Primary structure of the 'hinge' region of human IgG3. Probable
quadruplication of a 15-amino acid residue basic unit.";
J. Biol. Chem. 252:883-889(1977).
[3]
RN REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
RX MEDLINE=77021516; PubMed=823945;
Wolfenstein-Rodei C., Frangione B., Prelli F., Franklin E.C.;
"Structure of the PC fragment of immunoglobulin G3.";
Biochem. Biophys. Res. Commun. 71:907-914(1976).
[4]
RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
MEDLINE=82247835; PubMed=6808505;
Alexander A., Steinmetz M., Barritault D., Frangione B.,
Franklin E.C., Hood L., Buxbaum J.N.;
"Gamma Heavy chain disease in man: cDNA sequence supports partial
gene deletion model.";
Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -!- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
NORMALLY PRESENT IN THE HINGE REGION.
CC -!- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
REF.2.
CC -!- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
AND ALL OF THE CHI REGION.
CC -!- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
OF THE CHI REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
GAMMA-3 HEAVY CHAINS.
CC -!- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -!- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
SEGMENT (12-28).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
EMBL; J00231; AAA52805.1; ALT_SEQ.
HSSP; P01857; IFC1.
Genew; HGNC:5527;IGHG3.
MIM; 147120; -.
GO; GO:0005624; C:membrane fraction; NAS.
GO; GO:0003823; F:antigen binding activity; TAS.
GO; GO:0006955; P:immune response; NAS.
InterPro; IPR007110; Ig-like.
InterPro; IPR003597; Ig_C1.
Pfam; PF00047; Ig; 2.
SMART; SM00407; IgC1; 1.
PROSITE; PS00835; IG_LIKE; 2.
PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat;
Pyroliidone carboxylic acid.
FT DOMAIN 12 73 HINGE.
FT DOMAIN 74 183 CH2.
FT DOMAIN 184 289 CH3.
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 6 6 N-LINKED (GLCNAC...),
FT DISULFID 7 7 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 24 24 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 27 27 INTERCHAIN (WITH HEAVY CHAIN DIMER).

```

```

FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT CARBOHYD 140 140 N-LINKED (GLCNAC...).
FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
FT VARIANT 126 127 QV -> EB (IN ZUC).
FT VARIANT 134 134 P -> L (IN OMM).
FT VARIANT 139 139 F -> Y (IN OMM).
FT VARIANT 182 182 T -> A (IN OMM).
FT VARIANT 227 227 S -> N (IN OMM).
FT VARIANT 227 227 MISSING (IN ZUC).
FT VARIANT 279 279 F -> Y (IN OMM).
FT SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;

Query Match 39.5%; Score 1121.5; DB 1; Length 290;
Best Local Similarity 78.7%; Pred. No. 5.7e-73;
Matches 211; Conservative 11; Mismatches 21; Indels 25; Gaps 2;

QY 264 CDSKDFPVKYGKCPGLTGDGDTLDYYPEDTEGKVRHTCPCPAPEALGAPS 323
DB 48 CDTPPCPCPPEP-KSCDT-----PPCPCPAPELGGPS 82

QY 324 VLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKNVYDGVVHNNAKTKPREQYNST 383
DB 83 VLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVQKMYVDGVVHNNAKTKPREQYNST 142

QY 384 YRVSVLTVLHQDLNMGKCKYKCKSVNKALPVIEKTSKAKQPREPOVYTLPPPREMT 443
DB 143 FRVSVLTVLHQDLNMGKCKYKCKSVNKALPVIEKTSKAKQPREPOVYTLPPPREMT 202

QY 444 KQVSLTCLVKGFYPSDIAVEFESNQGNPNKYTPPPVLDSDGSFFLYSKLVDSKRWQO 503
DB 203 KQVSLTCLVKGFYPSDIAVEFESNQGNPNKYTPPPVLDSDGSFFLYSKLVDSKRWQO 262

QY 504 GNVFSCSVMEALHNNHYTKQSLSPGK 531
DB 263 GNVFSCSVMEALHNNHYTKQSLSPGK 290

RESULT 6
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-1 haplotype.";
RN Immunogenetics 18:387-397(1983).
RP SEQUENCE OF 1-128.

```

```

RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RN Biochem. J. 151:337-349(1975).
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RN Biochem. J. 116:249-259(1970).
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: M16426; AAA31289.1;
DR PIR; A91749; GHRE.
DR HSP; P01857; IPC1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00835; IG_LIKS; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 96 IG-LIKE 1.
FT DOMAIN 114 213 IG-LIKE 2.
FT DOMAIN 222 318 IG-LIKE 3.
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> B (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> B (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 32.4%; Score 921; DB 1; Length 323;
Best Local Similarity 61.1%; Pred. No. 1.3e-58;

```

	Matches	173;	Conservative	35;	Mismatches	45;	Indels	30;	Gaps	3;
Qy	252	VKAVT	SNVASVQCQNSDKFPVKYKPGKGCP	TGLGDEGDTLDLYDYYPEEDTEGDK-VRPHTC	310					
Dy	68	VWVTSSSQPVT	CNVA-----HPTATNTKVDKTVAPSTC	100						
Qy	311	--PCBPAEALGARSVELFPPPKD	TLMSRTPETCVVDVSHEDEVKNVWVGVEV	368						
Dy	101	SKPTCPPELGGGSVFIPPKD	TLMSRTPETCVVDVSQQDPPEVQFTWYNNEQV	160						
Qy	369	HNAXTKPREEQNSTYRVSVSLTVL	HDQLNGKEYKCKVSNKALFPVPIEKTI	428						
Dy	161	RTARPPLEREQFNSTIRVVSTLP	IPTHQDLWGKGFCKVHNKALPAPIEKTI	220						
Qy	429	EPOVYTLLPPSREEMTKNOVSLTC	LAVKGYPSDIAVEMESNGQPENNYKTITPV	488						
Dy	221	EPKYTWGPPEELSSRSVSUTCM	INGYPDSISVEWEKNGKAEDNYKITTTAV	280						
Qy	489	FLYSKLTVDKSRWQOGNVFCSVM	HEALHNHYTKQSLSLSPGK	531						
Dy	281	FYNKLSVPTSSEWORGDFVTC	SMHEALHNHYTKQSI	323						

## RESULT 7

GC2\_CAVPO  
ID GC2\_CAVPO  
STANDARD; PRT; 329 AA.

AC	PO1862;
AD	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	IG gamma-2 chain C region.
DE	IG gamma-2 chain C region.
OS	Cavia porcellus (Guinea pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX	NCBI_TaxID=10141;
ON	[1]
RP	SEQUENCE OF 1-3.
RA	Trischmann T.M.;
RL	Submitted (APR-1975) to the PIR data bank.
RN	[2]
RN	SEQUENCE OF 4-68.
RX	MEDLINE=71058471; PubMed=5538606;
RA	Birstein B.K., Hussain Q.Z., Cebra J.J.;
RT	"Structure of heavy chain from strain 13 guinea pig
RT	immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT	half-cystine joining heavy and light chains.";
RL	Biochemistry 10:18-25(1971).
RL	[3]
RN	

SEQUENCE OF 69-133 AND 312-329.  
MEDLINE=71058486; PubMed=5538616;  
Turner K.J., Cedra J.J.;  
"Structure of heavy chain from strain 13 guinea pig  
immunoglobulin-(2). II. Amino acid sequence of the carboxyl-terminal  
and hinge region cyanogen bromide fragments.";  
Biochemistry 10:9-17(1971).

[4]  
RN  
RP  
RX  
RA  
RT  
RT  
RT  
RL  
SEQUENCE OF 134-226.  
MEDLINE=75036072; Pubmed=4423665;  
Tracey D.E., Cebra J.G.;  
"Primary structure of the CH2 homology region from guinea pig IgG2  
antibodies.";  
Biochemistry 13:4796-4803(1974).

RN [5] PubMedEntry 25147756 400212577.  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; Pubmed=4609467;  
 RA Trischmann T.M., Gebra J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig I9G2  
 RT antibodies.";   
 RL Biochemistry 13:4804-4811 (1974).

RA Oliveira B., Lamm M.E.;  
RX MEDLINE=71058474; PubMed=4922544;  
RN DISULFIDE BONDS.  
[6]  
RA JACSchemistry 13:4004 4011(1974).

RT	"Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL	Biochemistry 10:26-31(1971).
RC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC	13 INBRED GUINEA PIGS.
DR	PIR; A94553; G2GP.
DR	HSSP; P01842; 7FAB.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003597; Ig cl.
DR	InterPro; IPR003006; IG_MHC.
DR	Pfam; PF00047; Ig: 2.
DR	SMART; SM00407; IGC1: 2.
DR	PROSITE; PS50835; IG LIKE; 3.
DR	PROSITE; PS00290; IG_MHC; 1.
KW	immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT	NON_TER 1
FT	DISULFID 16
FT	INTERCHAIN (WITH A LIGHT CHAIN) .
FT	DISULFID 28
FT	DISULFID 16
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
FT	DISULFID 105
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
FT	DISULFID 107
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
FT	DISULFID 110
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
FT	DISULFID 142
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
FT	CARBOHYD 178
FT	N-LINKED (GLCNAC. . . ) .
FT	DISULFID 178
FT	INTERCHAIN (WITH A HEAVY CHAIN) .
SC	SEQUENCE 329 AA: 36074 MW: 5D231B7164D1FBA9 CRC64:

```

Query Match      31.2%; Score 884.5; DB 1; Length 329;
Best Local Similarity 66.0%; Pred. No. 5.5e-56;
Matches 163; Conservative 28; Mismatches 45; Indels 11; Gaps 2

QY      295 YPEETEEDGKV-----RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPVTC 345
      :      :      :      :      :      :      :      :      :      :
Db      83 HPASTKVDKTVERTPTZPBPCPCPKPPENLGGPSVFIFPPKPKDTLMISLTPRVC 142

QY      346 VVVVDSHEDPEVKENWYDGYEVNATKREEQYNSTYRVSVLTVLQDMLNGKEYKC 405
      :      :      :      :      :      :      :      :      :      :
Db      143 VVVVDSQDEPVEQFTWFDVNNKPGVNAETKRVQYNTTFPRAVESVLPIQHDMLRGKEPKC 202

QY      406 KYSNKALPVIPIEKTISKAKGPREPPQVYTLPPSREEMTKNQVSLTCLVKGFPYPSDIAVEM 465
      :      :      :      :      :      :      :      :      :      :
Db      203 KVVNKALPAPIEKTISKTKGAPRPMDVYTLPPSRDELSSKSVSVTCLIIINFPFADIHVEM 262

QY      466 ENSGQP--ENNYKTPPVLDSGSGFFLYSKLTVYDKSRWQCNVPSCSVMHEALHNHYTK 523
      :      :      :      :      :      :      :      :      :      :
Db      263 ASNRVPVSEKSYKNTFPPIEDADGSGFFLYSKLTVYDKSAWDQGTVYTCVWHEALHNHYTK 322

QY      524 SLSLSPG 530
      :      :      :      :
Db      323 AISRSPG 329
      :      :      :      :

```

## RESULT 8

RESULT 8  
GC3\_MOUSE  
ID GC3\_MOUSE STANDARD; PRT; 329 AA.

AC P22436;  
DT 01-AUG-1991 (Rel. 19, Created)

DT	01-AUG-1991 (Rel. 19, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)

DE Ig gamma-3 chain C region, secreted form.  
OS *Mus musculus* (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Verte  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; M

```

OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.  
RX MEDLINE=85027161; PubMed=6092053;

RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P.  
RA Tucker P.W., Blattner F.R.;

RT "Structural analysis of the murine IgG3 constant  
RL EMBO J. 3:2041-2046(1984).  
CC

-----  
 CC This SWISS-PROT entry is copyright. It is provided  
 CC between the Swiss Institute of Bioinformatics  
 CC and the EMBL European Bioinformatics Centre  
 CC on the basis of a license agreement.

CC between the Swiss Institute of Bioinformatics  
CC the European Bioinformatics Institute. There

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; J00451; -; NOT\_ANNOTATED\_CDS.

DR PIR; B02156; G3MSC.

DR HSP; P01857; 1FC1.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.

DR InterPro; IPR003006; Ig\_MHC.

DR Pfam; PF00047; Ig; 3.

DR SMART; SM00407; IGc1; 2.

DR PROSITE; PS00835; IG\_LIKE; 3.

DR PROSITE; PS00290; IG\_MHC; 1.

KW Immunoglobulin domain, Immunoglobulin C region; Glycoprotein;

KW Transmembrane; Alternative splicing.

FT NON\_TER 1

FT DOMAIN 1 97

FT HINGE. 113

FT DOMAIN 114 223

FT CH2. 224

FT CH3. 224 327

SQ SEQUENCE 329 AA; 36228 MW; F45927174182BAD6 CRC64;

Query Match 29.6%; Score 841; DB 1; Length 329;

Best Local Similarity 66.2%; Pred. No. 7e-53;

Matches 151; Conservative 32; Mismatches 43; Indels 2; Gaps 1;

QY 306 RPTCP--PCPAPEALGAPSVLEPPKPKDLMISRTPEVTCVVDVSHDEPVKNVYV 363

DB 102 KPSTPFGSCPPGNILGGPSVFIFPPKPKDALMISLTPKVTVCVVDVSEDDPDVHVSFV 161

QY 364 DQVEVHNATKPREQYNSTYRWVSVLTVLHODWLNKGEYCKVSNKALPVPIETKISKA 423

DB 162 DNKEVHTANTQPREAQYNSTFRVSNALPIQHDWNRGKFEKCKVANKALPAPIETISRP 221

QY 424 KQPREPQVYTLPPREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 483

DB 222 KGRAQTPQVYTIPTPPREQMSKKVSLTCLVTNFFSEAISEVNERNGELEQDYKNTPPILD 281

QY 484 SGGSFLYSLKLTVDKSRWQGVFSCSVMEALHNHYTKSLSLSPGK 531

DB 282 SDGTFLYSLKLTVDTSWLQGEIFTCVSVHEALHNHHTQKLSRSPGK 329

RESULT 9

GC1\_RAT GC1\_RAT STANDARD; PRT; 326 AA.

AC P20759;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig gamma-1 chain C region.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89232738; PubMed=3149946;

RA Brueggemann M.;

RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";

RL Gene 74:473-482(1988).

DR PIR; PS0017; PS0017.

DR HSP; P01842; 7FAB.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig\_c1.

DR InterPro; IPR003006; Ig\_MHC.

DR Pfam; PF00047; Ig; 3.

DR SMART; SM00407; IGc1; 2.

DR PROSITE; PS00835; IG\_LIKE; 3.

DR PROSITE; PS00290; IG\_MHC; 1.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON\_TER 1 1

FT DOMAIN 1 97

FT HINGE. 112

FT DOMAIN 113 219

FT CH2. 220

FT CH3. 220 326

FT DISULFID 27 82

FT DISULFID 102 102

FT DISULFID 106 106

FT DISULFID 109 109

FT DISULFID 111 111

FT DISULFID 140 200

FT DISULFID 246 304

FT CARBOHYD 176 176

SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EP49B9DA CRC64;

Query Match 29.4%; Score 834.5; DB 1; Length 326;

Best Local Similarity 46.4%; Pred. No. 2e-52;

Matches 167; Conservative 47; Mismatches 73; Indels 73; Gaps 7;

QY 195 LOENSLYT---IPKGF-----GSHLLPFAFLHGNPWLNCCEILYFRRL 236

DB 17 LKNSMVTGLCLVKGYFPEPVTVWNSGALSSGVHFFPAVLQSG-----L 61

QY 237 QDNAENYVMKQVVDKAVTSNVASVQCDNSKDPVYKYPGKGCPTLGDGDTLDLYDP 296

DB 62 YTLTSSVTPESSWPSQTVCNVA-----HP 87

QY 297 EEDTEG-KVRPHT---CPCPAPEALGAPSVLEPPKPKDLMISRTPEVTCVVDVS 351

DB 88 ASSTKVDKKIIVRNCGGDKCKPCICTGS-EVSVFIFPPKPKDVLITLTPKVTCCVVVDIS 146

QY 352 HEDPEVFNMYVDGVEVHNATKPREQYNSTYRWVSVLTVLHODWLNKGEYCKVSNKA 411

DB 147 QDDPEVHFSFVDDVEVHTAQTPEPEQFNSTFRSVELPILHODWLNKGFTRCKVTSAA 206

QY 412 LPVPIETIKAKQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQP 471

DB 207 FPSPIETIKSGEPTQVPHVYTWSPTEKMTQNEVSIITCMWKGFPDPIYVEMQNGQP 266

QY 472 ENNYKTPPVLDSGFSFLYSLKLTVDKSRWQGVFSCSVMEALHNHYTKSLSLSPGK 531

DB 267 QENYKNTPTMTDGSYFLYSLKLVKVKERKQGGNTFTCSVLHGLHNHHTKLSLSPGK 326

RESULT 10

GC3M\_MOUSE

ID GC3M\_MOUSE STANDARD; PRT; 398 AA.

AC P03987;

DT 23-OCT-1986 (Rel. 02, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig gamma-3 chain C region, membrane-bound form.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85027161; PubMed=6092053;

RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,

RA Tucker P.W., Blattner F.R.;

RT "Structural analysis of the murine IgG3 constant region gene.";

RL EMBO J. 3:2041-2046(1984).

RN [2]

RP SEQUENCE OF 328-398 FROM N.A.

RX MEDLINE=84041483; PubMed=6314258;

RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,

RA Wall R.;

RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";

RL Nucleic Acids Res. 11:6775-6785(1983).

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; J00451; AAB59655.1; -;  
 DR EMBL; V01526; CAA24767.1; ALT\_SEQ.  
 DR PIR; A02156; G3MS.  
 DR HSSP; P01857; 1FC1.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 FT Transmembrane; Alternative splicing.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 FT TRANSMEM 346 362 POTENTIAL.  
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 333 333 E -> G (IN REF. 2).  
 FT CONFLICT 342 342 E -> Q (IN REF. 2).  
 FT CONFLICT 388 388 P -> F (IN REF. 2).  
 SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 29.2%; Score 830; DB 1; Length 398;  
 Best Local Similarity 65.9%; Pred. No. 5.5e-52;  
 Matches 149; Conservative 32; Mismatches 43; Indels 2; Gaps 1;  
 QY 306 RPTCP--PCPAEALGAPSVLPFPKPKDTLMISRTPEVTVVDVSHEDPEVKFNWTV 363  
 DB 102 KPSTPGSSCPPGNILGGSVFIFPKPKDLMISLTPKTVVDVSEDDPDVHVSFV 161  
 QY 364 DGEVHNATKPREQNSTYRVVSVLTVLHQLNGKEYCKVSKNKPVPFTIKTISKA 423  
 DB 162 DNKEVHTAWTPREAQYNSTFRVSVALPIQHQQDMRGKSEFKCKVNNKALPAPIERTISK 221  
 QY 424 KGQPREPOVTLPPSREEMTKQVSLTCLVKGYPSDIAVHESNGQPENNYKTPPVLD 483  
 DB 222 KGRAQTPQVYTIPTPREQMSKKVSLTCLVTFSEASISVEMERNGELQDQYKTPPILD 281  
 QY 484 SDGSFFLYSKLTVDKSRQCGNVFSCSVMEALHNNHYTOKSLSLSP 529  
 DB 282 SDGTFLYSLKLTVDTSKLGQGEFTCSVHVALHNNHTQNLRSRP 327

RESULT 11  
 GC1\_MOUSE ID GC1\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01568;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-1 chain C region secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RA "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 RT gamma 1 chain gene.";  
 RL Cell 18:559-568(1979).

[2] SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=8020559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 RT cloned in a bacterial plasmid.";  
 RL Gene 9:87-97(1980).  
 [3] SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salser W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 RT heavy chain.";  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 [4] SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Aduetubo K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 RT murine myeloma gammal chain.";  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 [5] DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
 RL Biochem. J. 126:837-850(1972).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=P01868-1; Sequence=Displayed;  
 CC Note=May be the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=P01869-1; Sequence=External;  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; V00793; CAA24172.1; -;  
 DR EMBL; V00793; CAA24173.1; -;  
 DR EMBL; V00793; CAA24174.1; -;  
 DR EMBL; V00793; CAA24175.1; -;  
 DR EMBL; V00795; CAA24176.1; -;  
 DR PIR; A02159; G1MS.  
 DR PDB; 1IGC; 03-JUN-95.  
 DR GlycoSuiteDB; P01868; -;  
 DR MGD; MGI:96446; Igh-4.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing; 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102  
 FT DISULFID 104 104  
 FT DISULFID 107 107  
 FT DISULFID 109 109  
 FT DISULFID 109 109 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).

```

FT DISULFID 138 198
FT CARBOHYD 174 174
FT FT N-LINKED (GLCNAC. . .)
/FTIG=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 29.1%; Score 826.5; DB 1; Length 324;
Best Local Similarity 46.2%; Pred. No. 7.5e-52; Indels 75; Gaps 8;
Matches 166; Conservative 50; Mismatches 68;

QY 196 QENSLYT---IPKGF-----GSHLLPFAFLHGNPWLNCNCEILYFRRLQ 237
DQ 18 QTNMVTGLCLVKGYFPEPVTVNWSGLSSGVHTFP-AVLQSD-----LY 62
QY 238 DNEAVNYWKQVVDVAVKATSVNASVQCDNSDKFPVKYPGKCPGLDGECDYLYYPE 297
DQ 63 TLSSSVTVFPSSPRPSETVTCNA-----HQA 88
QY 298 EDTEGDK---VRPHKCP--CPAPEALGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSH 352
DQ 89 SSTKVDKIVPRDCCKKPCICTVPE---VSSVFIAPPKPKDVLITLTPKVTVCVVVDISK 145
QY 353 EDPEVKFNWYDGVVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKKVSNNKAL 412
DQ 146 DQPEVQESFMFVDDVEVHTAQTQPREQNFSTFRSVELPIHQDLNGLKEYCKKVSNNAF 205
QY 413 VPTEKTSKAKQPREQVQVTLPPSRREMTNQVSLTCLVKGYFSPDAVWESNGOPE 472
DQ 206 PAPIEKTSKTRKPKAPQVVTIPPPKQMAKDKVSLTCDITDFPEDITVWQWNGQPA 265
QY 473 NNYKTPPVLDSDGSFFLYSLKLTVDKSRWQGNVFCSCVWHEALHNYTKSLSPGK 531
DQ 266 ENYKQTPQIMNINGSIFVTSKLVNKSWEAGNTFICSLVHGLHNNHHTKSLSPGK 324

RESULT 12
GCC_RAT ID_GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gere 74:473-482(1988).
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR; PS0018; PS0018.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1
FT DOMAIN 6 96 IG-LIKE 1.
FT DOMAIN 124 223 IG-LIKE 2.
FT DOMAIN 232 328 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80

```

```

FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 29.1%; Score 826.5; DB 1; Length 333;
Best Local Similarity 56.0%; Pred. No. 7.7e-52; Indels 27; Gaps 3;
Matches 158; Conservative 34; Mismatches 63;

QY 253 KAVTSNVA-SVQCDNSDKFPVKYFGK--CPTLGDGEGDTLDYYPEEDTEGDKVRPH 309
DQ 76 QTVTCNVAPASSTKVDKVERNGGIGHKPTC-----PT 111
QY 310 CPPCPAPEALGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVH 369
DQ 112 CHKCPVPELLGGPSVFIFPPKPKDILLISQNAKVTCTVVVDVSEEPDQVQFVFNVEVH 171
QY 370 NAKTKPREQYNSTYRVVSVLTVLHQDLNGLKEYCKKVSNNKALPVPTEKTSKAKQPRE 429
DQ 172 TAQTQPREQYNSTFRVVSALPIQHDMMSGKEFKCKVNNKALPSPTEKTSKPKGLVVK 231
QY 430 PQVYTLPPSRREMTNQVSLTCLVKGYFSPDAVWESNGOPENNYKTPPVLDSDGSFF 489
DQ 232 PQVYVWGPTEQLTEQTVSLTCLTSGFLPNDIGVWTSNGHIEKNYKNTPEVWMDSDGSFF 291
QY 490 LYSKLTVDKSRWQGNVFCSCVWHEALHNYTKSLSPGK 531
DQ 292 MYSKLVNRSWDSRAPFVFCVSVHGLHNNHHTKSLSPGK 333

RESULT 13
GCC_RAT ID_GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Bruggemann M.; Delmastro-Galfre P.; Waldmann H.; Calabi P.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X07189; CAA30169.1; -.
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1

```





```
QY 353 EDPEUKENWYDGVVHAKTKPREEQYNSTRVVSULTVLHODWLNKKEYKCVSNKAL 412
DB 146 DDPEVQFSFVDDVEVHTAQTPREEQNFSTRSVSELPFIHQDLWLNKKEFKCVNSAAF 205
QY 413 PVPICKTSKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGOPE 472
DB 206 PAPIKTSKTKGRPKAQVYTIIPPKQMAKDKUSLTCMTDFPEDIIVKQWNGQPA 265
QY 473 NNYKTPPVLSDGSEFFLYSKLTVDKSRWQOQNFVSCVMHEALHNHYTKSLSPG 530
DB 266 ENYKNTQPIMNTNGSYFYVSKLVQKNWEAGNTFTCSVLHGLHNHHTKSLSHSPG 323

RESULT 15
GCA_RAT STANDARD; PRT; 322 AA.
ID GCA_RAT
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2A chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13804; AAA41376.1; ALT_INIT.
CC PIR; PS0019; PS0019.
CC HSSP; P01842; 7FAB.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig_C1.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 2.
CC SMART; SM00407; IGc1; 2.
CC PROSITE; PS00835; IG LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1
FT DOMAIN 6 98 IG-LIKE 1.
FT DOMAIN 115 212 IG-LIKE 2.
FT DOMAIN 221 317 IG-LIKE 3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 28.8%; Score 817.5; DB 1; Length 322;
Best Local Similarity 46.3%; Pred. No. 3.2e-51;
Matches 165; Conservative 43; Mismatches 79; Indels 69; Gaps 6;

QY 195 LOENSLYT---IPKGF-----GSHLLPFAFLHGNPWLNCCEILYFRRWL 236
DB 17 LKNSMWTLGCLVKGYFPEPVTVTVNSGALSSGVHTFPAVLSG-----L 61
```

```
QY 237 QDAENVYWKOVVDKAVTNSVASVQCDNSDKPVPVYKPKGCPGLGDEGDTLDYYP 296
DB 62 YTLTSSVTVPSSTWSSQAVTCNVA-----HP 87
QY 297 EBDTEGD-KVRPHTCPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDP 355
DB 88 ASSTKVDKKIIVPRECNPCCGCTGS-EVSSVFIFPPKTKDVLITILTPKVTCCVVDISQNDP 146
QY 356 EVKFNWYDGVVEVHNAKTKPREEQYNSTRVVSULTVLHODWLNKKEYKCVSNKALPVP 415
DB 147 EVRFSWFIDDDVEVHTAQTHAPEKQSNSTLRSVSELPFVHODWLNKTKFKCVNSGAPPAP 206
QY 416 IEKTIKAKGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGOPENNY 475
DB 207 IEKSIKPEGTTPRGVGTMTAPFEEKETQSQSVITCMVKGFIPPDITYTEMKNGQPGQNY 266
QY 476 KTTTPVLSDGSEFFLYSKLTVDKSRWQOQNFVSCVMHEALHNHYTKSLSPG 531
DB 267 KNTPTMTDGSYFLYSLKLVKKETMQQGNFTCSVLHGLHNHHTKSLSHSPG 323
```

Search completed: October 14, 2003, 06:31:11  
Job time : 17.8186 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 55.8167 Seconds  
(without alignments)  
2454.925 Million cell updates/sec

Title: US-10-068-426-5  
Perfect score: 2839  
Sequence: 1 MPELLLLLLLPSLPHPIC.....MHEALHNYTKSLSLSPGK 531

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1616.5	56.9	626	Q8NG39	Q8ng39 homo sapien
2	1615.5	56.9	626	Q8NLF3	Q8nlf3 homo sapien
3	1208	42.6	701	Q96PQ8	Q96pq8 homo sapien
4	1205	42.4	471	Q8TC77	Q8tc77 homo sapien
5	1141	40.2	521	Q8N4Y9	Q8n4y9 homo sapien
6	1135	40.0	473	Q8TC63	Q8tc63 homo sapien
7	1130	39.8	509	Q8NF17	Q8nf17 homo sapien
8	1032	36.4	734	O35930	O35930 mus musculu
9	927	32.7	677	Q28256	Q28256 canis famli
10	891	31.4	337	Q95M34	Q95m34 equus cabal
11	842.5	29.7	469	Q8R3V9	Q8r3v9 mus musculu
12	839.5	29.6	463	Q99LC4	Q99lc4 mus musculu
13	835.5	29.4	437	Q9RLA4	Q9rla4 mus musculu
14	814.5	28.7	473	Q9D8L4	Q9d8l4 mus musculu
15	805	28.4	468	Q99L31	Q99l31 mus musculu
16	805	28.4	473	Q99L25	Q99l25 mus musculu

17	771	27.2	473	11	Q91Z05	Q91z05 mus musculu
18	771	27.2	474	11	Q8R3H6	Q8r3h6 mus musculu
19	359.5	12.7	375	4	Q9BSZ1	Q9bsz1 homo sapien
20	359.5	12.7	597	4	Q9BU10	Q9bu10 homo sapien
21	359.5	12.7	597	4	Q9BQB8	Q9bqb8 homo sapien
22	359.5	12.7	597	4	Q9BBB9	Q9bbb9 homo sapien
23	353	12.4	614	4	Q9GUA6	Q9gga6 homo sapien
24	349.5	12.3	588	4	Q8WUX4	Q8wux4 homo sapien
25	349.5	12.3	613	4	Q96EY0	Q96ey0 homo sapien
26	349.5	12.3	613	4	Q8WUK1	Q8wuk1 homo sapien
27	349.5	12.3	618	4	Q96AA6	Q96aa6 homo sapien
28	342	12.0	613	11	Q8VCX7	Q8vcx7 mus musculu
29	302.5	10.7	342	11	Q91XL1	Q91xl1 mus musculu
30	301	10.6	347	4	Q8N4F5	Q8n4f5 homo sapien
31	296	10.4	1328	5	Q21043	Q21043 caenorhabdi
32	284.5	10.0	486	11	Q91Z07	Q91z07 mus musculu
33	282.5	10.0	384	4	Q9UP60	Q9up60 homo sapien
34	282.5	10.0	493	4	Q8NCL6	Q8nc16 homo sapien
35	282.5	10.0	494	4	Q96K68	Q96k68 homo sapien
36	282.5	10.0	496	4	Q96KX8	Q96kx8 homo sapien
37	279.5	9.8	496	4	Q96DK0	Q96dk0 homo sapien
38	279.5	9.8	499	4	Q8NSK4	Q8nsk4 homo sapien
39	277.5	9.8	497	4	Q8WY24	Q8wy24 homo sapien
40	277.5	9.8	500	4	Q9BRV0	Q9brv0 homo sapien
41	276	9.7	557	4	Q8NSV4	Q8nsv4 homo sapien
42	274	9.7	416	4	Q9NPP6	Q9npp6 homo sapien
43	274	9.7	426	11	Q9DCD9	Q9dcd9 mus musculu
44	273	9.6	487	11	Q99KA4	Q99ka4 mus musculu
45	271.5	9.6	1527	5	Q9VZZ4	Q9vzz4 drosophila

ALIGNMENTS

RESULT 1

ID Q8NG39 PRELIMINARY; PRT; 626 AA.  
AC Q8NG39;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Platelet glycoprotein Ib alpha.  
GN GPIBA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Matsubara Y., Murata M., Sugita K., Ikeda Y.;  
RT "Identification of a novel point mutation in platelet glycoprotein  
RT Iba, Gly to Ser at residue 233, in a Japanese family with platelet-  
RT type von Willebrand disease.";  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB086948; BAC10305.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR Pfam; PF00560; LRR; 1.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR PRINTS; PR03217; PRICHEXTENS.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 5.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
SQ SEQUENCE 626 AA; 69985 MW; C7931FD07458B17F CRC64;

Query Match 56.9%; Score 1616.5; DB 4; Length 626;  
Best Local Similarity 84.2%; Pred. No. 2.1e-125;

Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy 1 MPELLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPDLPKDTTILHLENLLY 60  
 |||||  
 Db 1 MPELLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPDLPKDTTILHLENLLY 60  
 |||||

Qy 61 TFSLATMPYTRLTQNLDRCELTKLOVDGTLVGLGTLDSHQLQSLPLGQTLPALTV 120  
 |||||  
 Db 61 TFSLATMPYTRLTQNLDRCELTKLOVDGTLVGLGTLDSHQLQSLPLGQTLPALTV 120  
 |||||

Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 |||||  
 Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 |||||

Qy 181 AGLNGLENLDTLLQNSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRMWLODNA 240  
 |||||  
 Db 181 AGLNGLENLDTLLQNSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRMWLODNA 240  
 |||||

Qy 241 ENVVWKQVVDVKAITSNVASVQCDNSDKFPVVKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
 |||||  
 Db 241 ENVVWKQVVDVKAITSNVASVQCDNSDKFPVVKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
 |||||

Qy 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPPK-- 330  
 |||||  
 Db 301 EGDKVRATRTVVKFPTKAHTTPWGLFYWSSTASLDSQMPSSLHPTQESTKEQTTFPPRWT 360  
 |||||

Qy 331 PKDTL-----MISRTPEVT 344  
 |||||  
 Db 361 PNFTLHMSITFSKTPKST 379  
 |||||

RESULT 2

Q8N1F3 PRELIMINARY; PRT; 626 AA.

ID Q8N1F3;  
 AC Q8N1F3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glycoprotein Ib (Platelet), alpha polypeptide.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC027955; AAH27955.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_tyr.  
 DR InterPro; IPR002965; P\_Rich\_extensn.  
 DR Pfam; PF00560; LRR; 1.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 5.  
 DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
 SQ SEQUENCE 626 AA; 68969 MW; 1945PE52545DF573 CRC64;

Query Match 56.9%; Score 1615.5; DB 4; Length 626;  
 Best Local Similarity 84.2%; Pred. NO. 2.5e-125;  
 Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy 1 MPELLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPDLPKDTTILHLENLLY 60  
 |||||  
 Db 1 MPELLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPDLPKDTTILHLENLLY 60  
 |||||

Qy 61 TFSLATMPYTRLTQNLDRCELTKLOVDGTLVGLGTLDSHQLQSLPLGQTLPALTV 120  
 |||||  
 Db 61 TFSLATMPYTRLTQNLDRCELTKLOVDGTLVGLGTLDSHQLQSLPLGQTLPALTV 120  
 |||||

Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 |||||  
 Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 |||||

Qy 181 AGLNGLENLDTLLQNSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRMWLODNA 240  
 |||||  
 Db 181 AGLNGLENLDTLLQNSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRMWLODNA 240  
 |||||

Qy 241 ENVVWKQVVDVKAITSNVASVQCDNSDKFPVVKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
 |||||  
 Db 241 ENVVWKQVVDVKAITSNVASVQCDNSDKFPVVKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
 |||||

Qy 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPPK-- 330  
 |||||  
 Db 301 EGDKVRATRTVVKFPTKAHTTPWGLFYWSSTASLDSQMPSSLHPTQESTKEQTTFPPRWT 360  
 |||||

Qy 331 PKDTL-----MISRTPEVT 344  
 |||||  
 Db 361 PNFTLHMSITFSKTPKST 379  
 |||||

RESULT 3

Q96PQ8 PRELIMINARY; PRT; 701 AA.

ID Q96PQ8;  
 AC Q96PQ8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Factor VII active site mutant immunoconjugate.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=21477448; PubMed=11593034;  
 RA Hu Z., Gaten A.;  
 RT "Targeting tissue factor on tumor vascular endothelial cells and tumor  
 cells for immunotherapy in mouse models of prostatic cancer.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 DR EMBL; AF272774; AAK58686.1; -.  
 DR HSP; P00761; IAN1.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001438; EGF\_II.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR007110; I9-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig MHC.  
 DR InterPro; IPR001254; Ser. protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00047; Ig; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00010; EGFBLOOD.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00407; IGcl; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.

```

213 DB FESSLGTTQYICNVNHNKPSNTKVDKVKPEKSCDKTHTCPCPAPELLGGPSVFLFPKPK 277
QY 333 DTLMTSRTPETVCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTRYVVSFLTV 392
DB 273 DTLMTSRTPETVCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTRYVVSFLTV 332
QY 393 LHQDWLNGKEYKCKVSNKALPAPIEKTISAKAQGPKEPPQVYTLPPSRREMTKNQVSLTCL 452
DB 333 LHQDWLNGKEYKCKVSNKALPAPIEKTISAKAQGPKEPPQVYTLPPSRDELTKNQVSLTCL 392
QY 453 VKGFYPSDIAVEWSSNGQPENNYKTTPPVLDSDGSFFLSKLTVDKSRMQQGNVFSCSVM 512
DB 393 VKGFYPSDIAVEWSSNGQPENNYKTTPPVLDSDGSFFLSKLTVDKSRMQQGNVFSCSVM 452
QY 513 HEALHNHYTQKSLSLSPGK 531
DB 453 HEALHNHYTQKSLSLSPGK 471

RESULT 5
ID Q8N4Y9 PRELIMINARY; PRT; 521 AA.
AC Q8N4Y9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Primary B-Cells from Tonsils;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033178; AAH33178.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003357; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG_4.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
KW SEQUENCE 521 AA; 57156 MW; 2AC7D22E72D6CAA2 CRC64;

Query Match 40.2%; Score 1141; DB 4; Length 521;
Best Local Similarity 64.1%; Pred. No. 4.Se-86;
Matches 229; Conservative 18; Mismatches 56; Indels 54; Gaps 7

QY 194 LLOQNSLY-----TIKPGFGFGLHLLPAPLHGNPMLCNCEILYFRRLQDNAENNYVMK 247
DB 200 VLQSSGLSYLSSVTVFSSSLGT-----QTYTCN-----VNHKPSNTKVDK 240
QY 248 QVVDVKAVTSNVAV-----OCNDSKFPYKYPGK-----GPTLGDSDGDTLDYD 294
DB 241 R-VELTKPLGDTTHTCPRCPPEKSCDTPPPCPCEPKSCDTPPPCPCEPKSCD----- 295
QY 295 YPEEDTEGDKVRPHTCPCPAPEALGAPSVFLPPPKPDTLMTSRTPETVCVVVDVSHED 354
DB 296 -----TPPCPRCPAPELLGGPSVFLFPKPKDTLMISRTPETVCVVVDVSHED 344
QY 355 PEVKFNWYVDGVEVHNAKTKPREEQYNSTRYVVSFLTVLHQDWLNGKEYKCKVSNKALP 414
DB 345 PEVKFNWYVDGVEVHNAKTKPREEQYNSTRYVVSFLTVLHQDWLNGKEYKCKVSNKALP 404
QY 415 PIEKTISSAKAQGPKEPPQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWSSNGQPENN 474
DB 405 PIEKTISSAKAQGPKEPPQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWSSNGQPENN 464

```

```

QY 475 YKTPPVLDSDGFFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 YNTTPPMLDSDGFFFLYSKLTVDKSRWQGNVFCSCVMHEALHNRYTKSLSPGK 521

RESULT 6
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC025985; AH25985.1; -.
DR InterPro: IPR000923; BlueCu_1.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00196; COPPER_BLUE; 1.
DR PROSITE: PS00835; IG LIKE; 4.
DR PROSITE: PS00290; IG MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 40.0%; Score 1135; DB 4; Length 473;
Best Local Similarity 71.2%; Pred. No. 1.2e-85;
Matches 222; Conservative 17; Mismatches 40; Indels 33; Gaps 3;

QY 253 KAVTSNVASVQCNSDKFP-----VYKYPG-----KGCP 282
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
162 RSTSESTAALGCLVVKDYFFFPVTVSNWSGALTSGVHTFPAVLQSSGLYSLSVVTVFSS 221

QY 283 LGDEGTDLYDYYPEDTEGDKVRPH---TCPPCPAPEALGAPSVLPFPKPKDTLMISR 339
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 LGTKYTYCNVDHKPSNTKVDKRVESKYGPCCPAPFLGGPSVLPFPKPKDTLMISR 281

QY 340 TPEVTCVVVDVSHEDDEVKFNWVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLN 399
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 TPEVTCVVVDVSDQEDPEVQFNWVDGVEVHNKTKPREQFNSTYRVSVLTVLHODWLN 341

QY 400 GKEYKCKVSNKALPVPTEKTIISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPS 459
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
342 GKEYKCKVSNKGLPSSIEKTIISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPS 401

QY 460 DIAVESNGQPNNTYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 519
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 DIAVESNGQPNNTYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 461

QY 520 YTKSLSPGK 531
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 YTKSLSPGK 473

RESULT 7
Q8NF17 PRELIMINARY; PRT; 509 AA.
AC Q8NF17;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FLJ00385 protein (Fragment).
GN FLJ00385.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK090464; BAC03445.1; -.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig cl.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IGV; 3.
DR PROSITE: PS00835; IG LIKE; 3.
DR PROSITE: PS00290; IG MHC; 2.
FT NON TER 1
SQ SEQUENCE 509 AA; 56111 MW; 089498D8076E863C CRC64;

Query Match 39.8%; Score 1130; DB 4; Length 509;
Best Local Similarity 63.9%; Pred. No. 3.5e-85;
Matches 227; Conservative 18; Mismatches 56; Indels 54; Gaps 7;

QY 194 LLOENSLY-----TIPKFGFSGHLLPFAFLHGNPMLCNCEILYFRRLWLDNAENVVWK 247
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 VLQSSGLYSLSVVTVFSSSLGT-----QTYTCN-----VNHKPSNTKVDK 159

QY 248 QWVDKAVTSNVASV-----QCDNSDKFPVYKYPGK-----GCPTLGDEGTDLYD 294
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
160 R-VELKTPLGDTTHTCPRCPKPCSDTPPCPRCPKPCSDTPPCPRCPKPCSD---- 214

QY 295 YPEEDTEGDKVRPHTCPPCPAPEALGAPSVLPFPKPKDTLMISRTEVTCVVVDVSHED 354
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 -----TPPPCPCPAPEALGGSVLPFPKPKDTLMISRTEVTCVVVDVSHED 263

QY 355 PEVFNWVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPV 414
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 PEVQKMYVDGVEVHNKTKPREQFNSTYRVSVLTVLHODWLNKKEYCKVSNKALPV 323

QY 415 PIETISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVESNGQPNNT 474
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
324 PIETISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVESNGQPNNT 383

QY 475 YKTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSP 529
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 YNTTPPMLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNRYTKSLSP 438

RESULT 8
Q35930 PRELIMINARY; PRT; 734 AA.
AC Q35930;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Platelet glycoprotein Ib-alpha.
GN GPIBA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=98060158; PubMed=9410473;
RA Ware J., Russell S., Ruggeri Z.;
RT "Cloning of the murine platelet glycoprotein Ib(alpha) gene
RT highlighting species-specific platelet adhesion.";
RL Blood Cells Mol. Dis. 23:292-301(1997).
DB EMBL: U91967; AAC53320.1; -.

```



RT "Organization of the equine immunoglobulin heavy chain constant region  
 RI genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";  
 RL Immunobiology 199;105-119(1998).  
 DR EMBL; AJ300675; CAC44624.1; -;  
 DR InterPro; IPR007110; Ig\_Like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00407; IGH1; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 FT NON TER 1  
 SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFDIF6 CRC64;

Query Match 31.4%; Score 891; DB 6; Length 337;  
 Best Local Similarity 70.1%; Pred. No. 1.3e-65;  
 Matches 157; Conservative 34; Mismatches 31; Indels 2; Gaps 1;

QY 310 CPAPALGAPSVFLPPPKDITLMISRTPEVTCVVDVSHEDPEVKFNWYVQGVVH 369  
 DB 114 CPKAPALLGGPSVFIFPNPKDITLMITPEVTCVVDVSGENFDVKNWYMGVEVR 173  
 QY 370 NAKTPREOYNSTYRVSVLTVLHODWLNKGYKCKVSNKALPVPPIKTIKAKQPRE 429  
 DB 174 TATTPKEQFNSTYRVSVLRIQHODWLSGKFKCKVANNQALPQPIERTITTKGRSQE 233  
 QY 430 PQVTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQP--ENNYKTTTPVLDSGS 487  
 DB 234 PQVTLAPHDELKSKVSLTCLVKGFYPEINWQSNQGPPELEYKSTTQAQSDSGS 293  
 QY 488 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531  
 DB 294 YFLYSKSLVDNRWQQTFTTCGVMHEALHNHYTKQNVSKNPGK 337

## RESULT 11

Q8R3V9 PRELIMINARY; PRT; 469 AA.  
 ID Q8R3V9  
 AC Q8R3V9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical 52.0 kDa protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC024405; AAH24405.1; -;  
 DR InterPro; IPR007110; Ig\_Like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 469 AA; 51976 MW; 534793F15D05457 CRC64;

Query Match 29.7%; Score 842.5; DB 11; Length 469;  
 Best Local Similarity 47.1%; Pred. No. 2.2e-61;  
 Matches 169; Conservative 47; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLYT---IPKGF-----GSHLLPFAFHGNPWLNCIELYFRRLQ 237  
 DB 163 QTNQSVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFP-AVLQSD-----LY 207

QY 238 DNAENVYVWKQVVDVKAITSNVASQCDNSDKFPVKYKPGCGPTLGDGDTLDLYYPE 297  
 DB 208 TLSSSVTPSPSTWPSQTVCNVA-----HPA 227

Query Match 29.7%; Score 842.5; DB 11; Length 469;  
 Best Local Similarity 47.1%; Pred. No. 2.2e-61;  
 Matches 169; Conservative 47; Mismatches 68; Indels 75; Gaps 8;

QY 298 EDTGDK---VRPHTCPP--CPAPEALGAPSVFLPPPKDITLMISRTPEVTCVVDVSH 352  
 DB 234 SSTKVDKKIVPRDCGCKPCICTVPE---VSSVFIFPPPKDVLITITLPKVTCTVVDISK 290  
 QY 353 EDPEVKFNWYVQGVVHNAKTKPREOYNSTYRVSVLTVLHODWLNKGYKCKVSNKAL 412  
 DB 291 DDPEVQFSWFDVDEVHTAQTQPREQFNSTFRSVSELPIMHQDNLNGKFKCRVNSAAF 350  
 QY 413 PVPTEKTIKAKQPREPQVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 472  
 DB 351 PAPIEKTIKTKGRPKAPQVYTIIPPKKEQMAKQKVSITCMITDFFPEDITVEWQWNGQPA 410  
 QY 473 NNYKTTTPVLDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531  
 DB 411 ENYKNTQPIMDTDSYFVYSLNIVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 469

## RESULT 12

Q99LC4 PRELIMINARY; PRT; 463 AA.  
 ID Q99LC4  
 AC Q99LC4;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Similar to RIKEN cDNA 1810060009 gene.  
 GN IGH-4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC003435; AAH03435.1; -;  
 DR HSSP; P01842; 7FAB.  
 DR MGD; MGI:96446; IGH-4.  
 DR InterPro; IPR007110; Ig\_Like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 29.6%; Score 839.5; DB 11; Length 463;  
 Best Local Similarity 46.8%; Pred. No. 3.9e-61;  
 Matches 168; Conservative 48; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLYT---IPKGF-----GSHLLPFAFHGNPWLNCIELYFRRLQ 237  
 DB 157 QTNQSVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFP-AVLQSD-----LY 201

QY 238 DNAENVYVWKQVVDVKAITSNVASQCDNSDKFPVKYKPGCGPTLGDGDTLDLYYPE 297  
 DB 202 TLSSSVTPSPSTWPSQTVCNVA-----HPA 227

QY 298 EDTGDK---VRPHTCPP--CPAPEALGAPSVFLPPPKDITLMISRTPEVTCVVDVSH 352  
 DB 228 SSTKVDKKIVPRDCGCKPCICTVPE---VSSVFIFPPPKDVLITITLPKVTCTVVDISK 284

QY 353 EDPEVKFNWYVQGVVHNAKTKPREOYNSTYRVSVLTVLHODWLNKGYKCKVSNKAL 412  
 DB 285 DDPEVQFSWFDVDEVHTAQTQPREQFNSTFRSVSELPIMHQDNLNGKFKCRVNSAAF 344

QY 413 PVPTEKTIKAKQPREPQVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 472  
 DB 345 PAPIEKTIKTKGRPKAPQVYTIIPPKKEQMAKQKVSITCMITDFFPEDITVEWQWNGQPA 404

QY 473 NNYKTTTPVLDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531  
 DB 411 ENYKNTQPIMDTDSYFVYSLNIVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 469





Search completed: October 14, 2003, 06:19:32  
Job time : 58.8167 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 314.154 Seconds  
(without alignments)  
268.288 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPLLLLLLLLPSPHPHPIC.....MHEALHNHYTKSLSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2839	100.0	531	23	ABB78238
2	2829	99.6	531	23	ABB78237
3	2822.5	99.4	544	23	ABB78236
4	2822	99.4	531	23	ABB78239
5	2812.5	99.1	544	23	ABB78234
6	2812.5	99.1	544	23	ABB78235
7	2411.5	84.9	562	21	AAV49933
8	2381.5	83.9	568	21	AAV49935
9	1615.5	56.9	626	22	AAE12135

10	1536	54.1	250	23	ABB78244	Amino acid sequenc
11	1527	53.8	301	23	ABB78240	Amino acid sequenc
12	1526	53.8	290	23	ABB78243	Amino acid sequenc
13	1526	53.8	293	10	AA91368	45 kDa amino termi
14	1521.5	53.6	302	23	ABB78241	Amino acid sequenc
15	1521.5	53.6	610	17	AA89436	Mutated platelet g
16	1521.5	53.6	610	18	AAW18201	Platelet glycoprot
17	1520	53.5	290	23	ABB78245	Amino acid sequenc
18	1518.5	53.5	610	15	AA851116	Platelet glycoprot
19	1514.5	53.3	610	15	AA856664	Mutant platelet g1
20	1513.5	53.3	300	23	ABB78242	Amino acid sequenc
21	1237	43.6	313	19	AAW53321	P-selectin ligand
22	1237	43.6	313	20	AAW29766	P-selectin ligand
23	1237	43.6	313	23	AAW77946	Human dimeric P-se
24	1235.5	43.5	420	23	AAW15348	Human erythropoiet
25	1234	43.5	778	18	AAW15489	Pemphigus foliaceu
26	1230.5	43.3	401	18	AAW10537	Leptin 1-167/IGG1
27	1230	43.3	388	22	AAW50248	Human interleukin-
28	1230	43.3	449	20	AAW86278	Human noggin/immun
29	1230	43.3	502	24	ABJ37109	Concatameric immun
30	1229.5	43.3	542	23	AAU76357	Fc disintegrin con
31	1228.5	43.3	528	23	AAU79654	Human soluble P8G1
32	1226.5	43.2	397	18	AAW24060	Human obesity prot
33	1226.5	43.2	397	18	AAW22722	Human obesity prot
34	1226.5	43.2	399	18	AAW10536	Leptin 1-167/IGG1
35	1225.5	43.2	480	20	AAW90206	hB7.1Fc soluble fu
36	1225.5	43.2	480	24	ABU07263	Human expressed pr
37	1225.5	43.2	541	23	AAE29077	Human IL-22R-mutat
38	1225.5	43.2	558	23	AAE29076	Human IL-22R-mutat
39	1225.5	43.2	580	23	AAE23302	Human nectin-4-IgG
40	1225	43.1	476	22	AAW49243	Chimeric 4H6 anti-
41	1223	43.1	451	20	AAV50031	Human E27 anti-IgE
42	1223	43.1	451	20	AAW95859	Mus musculus anti-
43	1223	43.1	451	20	AAW95861	Mus musculus anti-
44	1223	43.1	451	20	AAW95663	Mus musculus anti-
45	1223	43.1	451	21	AAW07473	Amino acid sequenc

#### ALIGNMENTS

#### RESULT 1

ABB78238	
ID	ABB78238 standard; Protein; 531 AA.
XX	
AC	ABB78238;
XX	
DT	25-NOV-2002 (first entry)
XX	
DE	Amino acid sequence of GPIIb290/2V-Ig fusion protein.
XX	
KW	Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW	leukocyte; platelet activation; ischemic heart disease; atherosclerosis;
KW	acute myocardial infarction; stroke; venous thrombosis; thrombosis;
KW	arterial thrombosis; angina; vascular condition; vascular inflammation;
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 487 /note= "Ser encoded by CCC"
FT	
XX	
PN	WO200263003-A2.
XX	
PD	15-AUG-2002.
XX	
PF	06-FEB-2002; 2002WO-US03549.
XX	
PR	06-FEB-2001; 2001US-266838P.
XX	
PA	(GEMV ) GENETICS INST LLC.
XX	

PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
DR N-PSDB; ABQ78665.  
XX  
PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
PS Claim 20; Page 3-4; 45pp; English.  
XX  
CC The present sequence represents a fusion protein of glycoprotein  
CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX  
SQ Sequence 531 AA;  
  
Query Match 100.0%; Score 2839; DB 23; Length 531;  
Best Local Similarity 100.0%; Pred. No. 7.9e-204;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MPLLILLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
Db 1 MPLLILLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
QY 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQQLSLPLGOTLPALT 120  
Db 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQQLSLPLGOTLPALT 120  
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
QY 241 SKAGQPREQVYTLPPSRQEMTKNQVSLTCLVKGFPDSQIAVWESNGQPPENNYKTTP 480  
Db 241 SKAGQPREQVYTLPPSRQEMTKNQVSLTCLVKGFPDSQIAVWESNGQPPENNYKTTP 480  
QY 481 VLDSDGSFFLYSKLTVDKSRQCGNVFSCSVMHEALHNNHYTKLSLSLSPGK 531  
Db 481 VLDSDGSFFLYSKLTVDKSRQCGNVFSCSVMHEALHNNHYTKLSLSLSPGK 531  
  
RESULT 2  
ABB78237  
ID ABB78237 standard; Protein; 531 AA.  
XX  
AC ABB78237;  
XX  
DT 25-NOV-2002 (first entry)  
XX  
DE Amino acid sequence of GPIb290-Ig fusion protein.

XX Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 487 /note= "Ser encoded by CCC"  
XX  
PN WO200263003-A2.  
XX 15-AUG-2002.  
XX 06-FEB-2002; 2002WO-US03549.  
XX 06-FEB-2001; 2001US-266838P.  
XX (GEMY ) GENETICS INST LLC.  
XX  
PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
DR N-PSDB; ABQ78664.  
XX  
PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
PS Claim 20; Page 3; 45pp; English.  
XX  
CC The present sequence represents a fusion protein of glycoprotein  
CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX  
SQ Sequence 531 AA;  
  
Query Match 99.6%; Score 2829; DB 23; Length 531;  
Best Local Similarity 99.6%; Pred. No. 4.4e-203;  
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MPLLILLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
Db 1 MPLLILLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
QY 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQQLSLPLGOTLPALT 120  
Db 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQQLSLPLGOTLPALT 120  
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
QY 241 ENVYVWKQVVDVKA VTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYVWKQVVDVKA VTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300  
QY 301 EGDVVRPHTCPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
Db 301 EGDVVRPHTCPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360

QY 361 WYVDGVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGYCKVSNKALPVPPIEKTI 420  
DB 361 WYVDGVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGYCKVSNKALPVPPIEKTI 420  
QY 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPP 480  
DB 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPP 480  
QY 481 VLSDSGSFLLYKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSLSPGK 531  
DB 481 VLSDSGSFLLYKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSLSPGK 531  
RESULT 3  
ID ABB78236  
XX ABB78236 standard; Protein; 544 AA.  
AC ABB78236;  
DT 25-NOV-2002 (first entry)  
DE Amino acid sequence of GPlb302/4X-Ig fusion protein.  
KW Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Misc-difference 500 /note= "Ser encoded by CCC"  
FT  
XX WO200263003-A2.  
PN 15-AUG-2002.  
PD 06-FEB-2002; 2002WO-US03549.  
PF 06-FEB-2001; 2001US-266838P.  
PR (GEMV ) GENETICS INST LLC.  
PA Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
PI WPI; 2002-657537/70.  
DR N-PSDB; ABQ78663.  
XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX Claim 20; Page 3; 45pp; English.  
XX The present sequence represents a fusion protein of glycoprotein  
CC 1B-alpha (GPlb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX Sequence 544 AA;  
Query Match 99.4%; Score 2822.5; DB 23; Length 544;  
Best Local Similarity 97.6%; Pred. No. 1.4e-202; Indels 13; Gaps 1;  
Matches 531; Conservative 0; Mismatches 0;

QY 1 MPLLLLLLLPSPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDITLHLENLLY 60  
DB 1 MPLLLLLLLPSPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDITLHLENLLY 60  
QY 61 TFSLATLMPYTRITQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLLGOTLPALT 120  
DB 61 TFSLATLMPYTRITQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLLGOTLPALT 120  
QY 121 LDVSENLRLSLPGLARGELGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
DB 121 LDVSENLRLSLPGLARGELGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLP 180  
QY 181 AGLLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240  
DB 181 AGLLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240  
QY 241 ENYVWKQVVDVKAVTSNVASVQCDNSDKPPVYKYPGKGCPTLGDGEGDTLDLYYPEEDT 300  
DB 241 ENYVWKQVVDVKAVTSNVASVQCDNSDKPPVYKYPGKGCPTLGDGEGDTLDLYYPEEDT 300  
QY 301 EGDKV-----RPHCTCPCPAPEALGAPSVFLFPKPKDITLMSRTPVTCV 347  
DB 301 EGDKVAATATVVKFPTKARPHCTCPAPEALGAPSVFLFPKPKDITLMSRTPVTCV 360  
QY 348 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGYCKV 407  
DB 361 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGYCKV 420  
QY 408 SNKALPVPPIEKTISKAGOPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVWES 467  
DB 421 SNKALPVPPIEKTISKAGOPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVWES 480  
QY 468 NGQPENNYKTTPPVLDSDSGSFLLYKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSL 527  
DB 481 NGQPENNYKTTPPVLDSDSGSFLLYKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSL 540  
QY 528 SPGK 531  
DB 541 SPGK 544  
RESULT 4  
ID ABB78239  
XX ABB78239 standard; Protein; 531 AA.  
AC ABB78239;  
DT 25-NOV-2002 (first entry)  
DE Amino acid sequence of GPlb290/1A-Ig fusion protein.  
KW Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Misc-difference 253 /note= "Val encoded by GCG"  
FT Misc-difference 487 /note= "Ser encoded by CCC"  
XX WO200263003-A2.  
XX 15-AUG-2002.  
XX 06-FEB-2002; 2002WO-US03549.  
XX 06-FEB-2001; 2001US-266838P.

(GEMY ) GENETICS INST LLC.

Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 WPI: 2002-657537/70.  
 N-PSDB; ABQ78666.

New glycoprotein Ib alpha fusion polypeptides, useful for treating a disorder associated with platelet activation e.g. ischaemic heart disease, stroke, venous or arterial thrombosis or atherosclerosis -

Claim 20; Page 4; 45pp; English.

The present sequence represents a fusion protein of glycoprotein Ib-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion protein inhibits the adherence of platelets to leukocytes. The fusion polypeptide is useful for treating a disorder associated with platelet activation e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It can also be used to treat vascular conditions associated with vascular inflammation, thrombosis, and angioplasty-related restenosis.

Sequence 531 AA;

Query Match 99.4%; Score 2822; DB 23; Length 531;  
 Best Local Similarity 99.4%; Pred. No. 1.5e-202;  
 Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLLLLLLLLLSPHPHPICVSVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
 DB 1 MLLLLLLLLLSPHPHPICVSVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
 QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLVLGTLDSLHNLQSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLVLGTLDSLHNLQSLPLGOTLPALT 120  
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
 DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
 QY 181 AGLNGLENLDTLLQENSLYTIKPGFSGHLLPFAFHGNPMLCNCEILYFRRLQDNA 240  
 DB 181 AGLNGLENLDTLLQENSLYTIKPGFSGHLLPFAFHGNPMLCNCEILYFRRLQDNA 240  
 QY 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300  
 DB 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300  
 QY 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
 DB 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
 QY 361 WYVDGVEVHNKTKPREEQYNSTRVSVLTVLQDHALNGKEYCKYKSNKALPVP-EKTI 420  
 DB 361 WYVDGVEVHNKTKPREEQYNSTRVSVLTVLQDHALNGKEYCKYKSNKALPVP-EKTI 420  
 QY 421 SKAKQPREPQVYTLPPSREEMTKNOVSLTCLVKGFPSPDI AVEWESNGQPENNYKTTP 480  
 DB 421 SKAKQPREPQVYTLPPSREEMTKNOVSLTCLVKGFPSPDI AVEWESNGQPENNYKTTP 480  
 QY 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVWHEALHNHYTKLSLSPGK 531  
 DB 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVWHEALHNHYTKLSLSPGK 531

RESULT 5  
 ABB78234  
 ID ABB78234 standard; Protein; 544 AA.  
 XX  
 AC ABB78234;  
 XX  
 DT 25-NOV-2002 (first entry)

XX Amino acid sequence of GPIb302-Ig fusion protein.  
 DE Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 XX leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 500 /note= "Ser encoded by CCC"  
 FT  
 XX WO200263003-A2.  
 XX 15-AUG-2002.  
 XX 06-FEB-2002; 2002WO-US03549.  
 XX 06-FEB-2001; 2001US-266838P.  
 XX (GEMY ) GENETICS INST LLC.  
 XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX WPI: 2002-657537/70.  
 XX N-PSDB; ABQ78661.  
 XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a disorder associated with platelet activation e.g. ischaemic heart disease, stroke, venous or arterial thrombosis or atherosclerosis -

Claim 5; Page 2; 45pp; English.

The present sequence represents a fusion protein of glycoprotein Ib-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion protein inhibits the adherence of platelets to leukocytes. The fusion polypeptide is useful for treating a disorder associated with platelet activation e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It can also be used to treat vascular conditions associated with vascular inflammation, thrombosis, and angioplasty-related restenosis.

Sequence 544 AA;

Query Match 99.1%; Score 2812.5; DB 23; Length 544;  
 Best Local Similarity 97.2%; Pred. No. 7.9e-202;  
 Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

QY 1 MLLLLLLLLLSPHPHPICVSVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
 DB 1 MLLLLLLLLLSPHPHPICVSVASHLEVNCNRLTALPPDLPKDTTILHLENLLY 60  
 QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLVLGTLDSLHNLQSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLVLGTLDSLHNLQSLPLGOTLPALT 120  
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
 DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
 QY 181 AGLNGLENLDTLLQENSLYTIKPGFSGHLLPFAFHGNPMLCNCEILYFRRLQDNA 240  
 DB 181 AGLNGLENLDTLLQENSLYTIKPGFSGHLLPFAFHGNPMLCNCEILYFRRLQDNA 240  
 QY 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300  
 DB 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300  
 QY 301 EGDKV-----RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347

```

|||||
301 EGDKYRATRTVVKFPTKARHTCCPCPAPEALGAPSVLFPPKPKDTLMISRTPEVTCVV 360
348 VDVSHEDEPEVKFNMYVDGVEVHNAKTKPREEQYNSTVRVSVLTVLHQDWLNGKEYKCKV 407
361 VDVSHEDEPEVKFNMYVDGVEVHNAKTKPREEQYNSTVRVSVLTVLHQDWLNGKEYKCKV 420
408 SNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 467
421 SNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 480
468 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVPCFSCSMHEALHNHYTQKSLSL 527
481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVPCFSCSMHEALHNHYTQKSLSL 540
528 SPGK 531
541 SPGK 544

RESULT 6
ABB78235
ID ABB78235 standard; Protein; 544 AA.
XX AC
XX ABB78235;
XX DT
XX 25-NOV-2002 (first entry)
XX DE
XX Amino acid sequence of GPIb302/2A-Ig fusion protein.
XX KW
XX Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX KW leukocyte; platelet activation; ischaemic heart disease;
XX KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX KW arterial thrombosis; angina; vascular condition; vascular inflammation;
XX KW thrombosis; angioplasty; restenosis.
XX OS
XX Synthetic.
XX PH
XX Key Location/Qualifiers
XX FT Misc-difference 500
XX FT /note= "Ser encoded by CCC"
XX PN
XX WO200263003-A2.
XX PD
XX 15-AUG-2002.
XX PF
XX 06-FEB-2002; 2002WO-US03549.
XX PR
XX 06-FEB-2001; 2001US-266838P.
XX PA (GEMY ) GENETICS INST LLC.
XX PI
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX DR
XX WPI; 2002-657537/70.
XX DR N-PSDB; ABQ78662.
XX PT
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX PT disorder associated with platelet activation e.g. ischaemic heart
XX PT disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX PS
XX Claim 20; Page 3; 45pp; English.
XX CC
XX The present sequence represents a fusion protein of glycoprotein
XX CC 1b-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion
XX CC protein inhibits the adherence of platelets to leukocytes. The fusion
XX CC polypeptide is useful for treating a disorder associated with platelet
XX CC activation e.g. ischaemic heart disease, acute myocardial infarction,
XX CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or
XX CC unstable angina. It can also be used to treat vascular conditions
XX CC associated with vascular inflammation, thrombosis, and
XX CC angioplasty-related restenosis.

```

```

SQ Sequence 544 AA;
Query Match 99.1%; Score 2812.5; DB 23; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.9e-202;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

QY 1 MPLLALLLLPSPLPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
DB 1 MPLLALLLLPSPLPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
QY 61 TFSLATLMPYTRLTQLNDRCELTKLQVDTGLPVLGTLDSLHQLQSLPLIGOTLPALT 120
DB 61 TFSLATLMPYTRLTQLNDRCELTKLQVDTGLPVLGTLDSLHQLQSLPLIGOTLPALT 120
QY 121 LDVSFNRLTSLPLCALRGELQELYLKGNELKTLPPGLLTPTPKLKLKSLANNLTLP 180
DB 121 LDVSFNRLTSLPLCALRGELQELYLKGNELKTLPPGLLTPTPKLKLKSLANNLTLP 180
QY 181 AGLLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240
DB 181 AGLLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240
QY 241 ENVYVWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCTPLGDEGDTLDLYYPEEDT 300
DB 241 ENVYVWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCTPLGDEGDTLDLYYPEEDT 300
QY 301 EGDKV-----RPHTCPPCPAPEALGAPSVLFPPKPKDTLMISRTPEVTCVV 347
DB 301 EGDKVAATATVVKFPTKARHTCCPCPAPEALGAPSVLFPPKPKDTLMISRTPEVTCVV 360
QY 348 VDVSHEDEPEVKFNMYVDGVEVHNAKTKPREEQYNSTVRVSVLTVLHQDWLNGKEYKCKV 407
DB 361 VDVSHEDEPEVKFNMYVDGVEVHNAKTKPREEQYNSTVRVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 467
DB 421 SNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 480
QY 468 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVPCFSCSMHEALHNHYTQKSLSL 527
DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVPCFSCSMHEALHNHYTQKSLSL 540
QY 528 SPGK 531
DB 541 SPGK 544

RESULT 7
AAV49933
ID AAV49933 standard; Protein; 562 AA.
XX AC
XX AAV49933;
XX DT
XX 01-FEB-2000 (first entry)
XX DE
XX Human glycoprotein Ib/mouse IgG1Fc chimeric protein.
XX KW
XX Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;
XX KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
XX KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX OS
XX Chimeric - Homo sapiens.
XX OS Chimeric - Mus musculus.
XX FH
XX Key Location/Qualifiers
XX FT Peptide 1..16
XX FT /label= signal
XX FT Protein 17..562
XX FT /label= human glycoprotein_Ib/mouse_IgG1Fc_chimeric
XX FT protein
XX PN
XX WO9954360-A1.
XX

```

```
PD 28-OCT-1999.
PF 13-JAN-1999; 99WO-JP00089.
PR 23-APR-1998; 98JP-0113962.
PA (AJIN ) AJINOMOTO CO INC.
PI Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;
PI Tanaka A, Shinozaki J, Jojima Y;
DR WPI; 2000-013233/01.
DR N-PSDB; AAZ35701.
XX
XX Direct quantitative detection of glycoalkaloid with immobilized von
PT Willebrand factor to bond with chimeric protein via inhibiting
PT glycoprotein Ib binding, for diagnosis of thrombotic diseases and
PT screening anti-thrombotic substances -
XX
XX Example 1; Page 70-71; 83pp; Japanese.
XX
XX A method has been developed for the detection of binding between the von
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in
CC which the von Willebrand factor immobilised in a reactor reacts with
CC glycoprotein Ib in the presence of a binding inducer to promote binding
CC between the von Willebrand factor and glycoprotein Ib. This method is
CC for the detection of glycoalkaloid as a means of thrombotic disease
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also
CC for screening substances with anti-thrombotic activity for the
CC prevention and treatment of thrombotic diseases. The method is direct,
CC convenient and quantitative, with reproducibility, and there is no need
CC to construct a monoclonal antibody for the assay. The present sequence
CC is a human glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric
CC protein from the present invention.
XX
XX SQ Sequence 562 AA;
Query Match 84.9%; Score 2411.5; DB 21; Length 562;
Best Local Similarity 79.5%; Pred. No. 8e-172;
Matches 449; Conservative 40; Mismatches 39; Indels 37; Gaps 4;
QY 1 MPELLLLLLLSPHPHPCVSVASHLEVNCNRLTALPPDLPKDTILHLSENLLY 60
DB 1 MPELLLLLLLSPHPHPCVSVASHLEVNCNRLTALPPDLPKDTILHLSENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHNQLSLPLGQTLPALTV 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHNQLSLPLGQTLPALTV 120
QY 121 LDVSFNRLTSLPLGALRGELGLOELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTPEL 180
DB 121 LDVSFNRLTSLPLGALRGELGLOELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTPEL 180
QY 181 AGLINGLENLDTLLLOENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEIILYFRWLQDNA 240
DB 181 AGLINGLENLDTLLLOENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEIILYFRWLQDNA 240
QY 241 ENVYVWKQVDVAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYPEEDT 300
DB 241 ENVYVWKQVDVAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYPEEDT 300
QY 301 EGDQKVR-----PHT-----CPP--CPAPEALGAPSVFL 326
DB 301 EGDQKVRATRTVVKPPTKARTTPWGLFYSWSTASLDVPRDCCKPCICTVPE---VSSVFI 357
QY 327 FPPKPKDTLMSRTEVTCVVVDVSHEDPEVKFVYVGVGVHNAKTKPREQYNSTYRV 386
DB 327 FPPKPKDTLMSRTEVTCVVVDVSHEDPEVKFVYVGVGVHNAKTKPREQYNSTYRV 386
QY 387 VSVLTVLHODMLNGKEYCKYKSNKALPVPIEKTISKAKGQPREPVVTLPPSREEMTKNQ 446
DB 387 VSVLTVLHODMLNGKEYCKYKSNKALPVPIEKTISKAKGQPREPVVTLPPSREEMTKNQ 446
QY 418 VSELPIMHQDMLNGKEFKRNSAFAFPAPIBKTIISKTKGRPKAPQVITIIPPKQEQMAOK 477
DB 418 VSELPIMHQDMLNGKEFKRNSAFAFPAPIBKTIISKTKGRPKAPQVITIIPPKQEQMAOK 477
```

```
QY 447 VSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 506
DB 478 VSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 537
QY 507 FSCSVWHEALHNHYTQKSLSLSPGK 531
DB 538 FSCSVWHEALHNHYTQKSLSLSPGK 562
XX
XX RESULT 8
XX ID AAY49935 standard; Protein; 568 AA.
XX AC AAY49935;
XX DT 01-FEB-2000 (first entry)
XX DE Human glycoprotein Ib/mouse IgG2aFc chimeric protein #2.
XX KW Glycoprotein Ib; glycoalkaloid; detection; antithrombotic; binding;
XX KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
XX KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Mus musculus.
XX PH Key Location/Qualifiers
XX FT Peptide 1..16
XX FT Protein 17..568
XX FT /label= signal
XX FT /label= Human glycoprotein_Ib/mouse_IgG2aFc_chimeric
XX PN WO9954360-A1.
XX PD 28-OCT-1999.
XX PF 13-JAN-1999; 99WO-JP00089.
XX PR 23-APR-1998; 98JP-0113962.
XX PA (AJIN ) AJINOMOTO CO INC.
XX PI Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;
XX PI Tanaka A, Shinozaki J, Jojima Y;
XX DR WPI; 2000-013233/01.
XX DR N-PSDB; AAZ35706.
XX
XX Direct quantitative detection of glycoalkaloid with immobilized von
PT Willebrand factor to bond with chimeric protein via inhibiting
PT glycoprotein Ib binding, for diagnosis of thrombotic diseases and
PT screening anti-thrombotic substances -
XX
XX Example 1; Page 77-79; 83pp; Japanese.
XX
XX A method has been developed for the detection of binding between the von
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in
CC which the von Willebrand factor immobilised in a reactor reacts with
CC glycoprotein Ib in the presence of a binding inducer to promote binding
CC between the von Willebrand factor and glycoprotein Ib. This method is
CC for the detection of glycoalkaloid as a means of thrombotic disease
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also
CC for screening substances with anti-thrombotic activity for the
CC prevention and treatment of thrombotic diseases. The method is direct,
CC convenient and quantitative, with reproducibility, and there is no need
CC to construct a monoclonal antibody for the assay. The present sequence
CC represents a human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc
CC chimeric protein from the present invention.
XX
XX SQ Sequence 568 AA;
```

Query Match

83.9%; Score 2381.5; DB 21; Length 568;



```

Best Local Similarity 79.4%; Pred. No. 1.4e-169;
Matches 451; Conservative 29; Mismatches 51; Indels 37; Gaps 3;

QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLENLY 60
   |||||
Db 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLENLY 60

QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQSLPLLGOTLPALT 120
   |||||
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQSLPLLGOTLPALT 120

QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLT 180
   |||||
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLT 180

QY 181 AGLNGLENLDTLLQENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELLYFRRL 240
   |||||
Db 181 AGLNGLENLDTLLQENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELLYFRRL 240

QY 241 ENYVWKQVVDKAVTSNVASVQCDNSKFPVYKPGKCPGLGDEGTDLYDY 300
   |||||
Db 241 ENYVWKQVVDKAVTSNVASVQCDNSKFPVYKPGKCPGLGDEGTDLYDY 300

QY 301 EGDKVR-----PHT-----CPA---CRAPEALGAPS 323
   |||||
Db 301 EGDKVRATRVVKFPTKAHTTPWGLFYSWSTASLDEPRGTIKPCPCPKCAPN 360

QY 324 VLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVAFNVDGVEVHNAKTPRE 383
   |||||
Db 361 VLFPPKPKDVLMSISPIVTCVVDVSEDDPQVQISFVNVEVHTAQTHRD 420

QY 384 YRVSVVLTVHQDLNGKYEKCKVSNKALPVFIETISKAKOPREPQVYTLPP 443
   |||||
Db 421 LRVVSALPIQHQDWSGKEFKCKVANKOLPAPIERTISKPKGSVRAPQVY 480

QY 444 KQVSLTCLVKGFYPSDIAVESKCPENNYKTPVLDSDGSFSLYKLTVXS 503
   |||||
Db 481 KQGVLTTCMTDFMPEDIYVETWNGKTELANKTEPVLDSGYSFYMKLVR 540

QY 504 GNVFSCVNHAEALHNYTOKSLSPGK 531
   |||||
Db 541 RNSYSCSVVHEGLHNNHTKFSRTPGK 568

RESULT 9
AAE12135
ID AAE12135 standard; Protein; 626 AA.
AC AAE12135;
XX
DT 03-JAN-2002 (first entry)
XX
DE Human glycoprotein Ib (platelet) alpha (GP1BA).
KW Human; haplotyping; glycoprotein Ib (platelet) alpha protein; GP1BA;
KW Bernard-Soulier syndrome; platelet-type von Willebrand disease; HIV;
KW Alzheimer's disease; human immunodeficiency virus; SNP;
KW single nucleotide polymorphism; chromosome 17pter-p12.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 72
   /note= "Arg at this position is replaced with His
   due to single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 86
   /note= "Leu at this position is replaced with Phe
   due to single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 161
   /note= "Thr at this position is replaced with Met
   due to single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 592
   /note= "Arg at this position is replaced with His

```

```

FT Misc-difference 624
   /note= "His at this position is replaced with Arg
   due to single nucleotide polymorphism (SNP)"
XX
PN WO200175065-A2.
XX
XX 11-OCT-2001.
XX
XX 03-APR-2001; 2001WO-US10671.
XX
XX 03-APR-2000; 2000US-194341P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Bentivegna SC, Choi JY, Kliem SE, Koshy B, Parks KE;
XX
XX WPI; 2001-626427/72.
XX
XX N-PSDB; AAD20681.
XX
XX New haplotypes of the glycoprotein Ib platelet alpha polypeptide gene
   are useful for diagnosis and drug discovery for treating Bernard
   Soulier syndrome, platelet-type von Willebrand disease, HIV and
   Alzheimer's disease
XX
XX Claim 29; Fig 3; 66pp; English.
XX
XX The invention relates to methods for haplotyping glycoprotein Ib
   (platelet) alpha polypeptide (GP1BA) gene of an individual. The
   method involves determining if the individual has one of the GP1BA
   haplotypes or haplotype pairs. The methods of the invention are
   useful for disease diagnosis and in the discovery and development
   of drugs for treating diseases associated with GP1BA activity e.g.
   Bernard-Soulier syndrome, platelet-type von Willebrand disease, HIV
   and Alzheimer's disease. The present sequence is human GP1BA protein.
   GP1BA gene is located on chromosome 17pter-p12.
XX
XX Sequence 626 AA;
XX
XX Query Match 56.9%; Score 1615.5; DB 22; Length 626;
   Best Local Similarity 84.2%; Pred. No. 3.2e-112;
   Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLENLY 60
   |||||
Db 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLENLY 60

QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQSLPLLGOTLPALT 120
   |||||
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQSLPLLGOTLPALT 120

QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLT 180
   |||||
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLT 180

QY 181 AGLNGLENLDTLLQENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELLYFRRL 240
   |||||
Db 181 AGLNGLENLDTLLQENSLYTIKGFSGSHLLPFAFLHGNPWLNCCELLYFRRL 240

QY 241 ENYVWKQVVDKAVTSNVASVQCDNSKFPVYKPGKCPGLGDEGTDLYDY 300
   |||||
Db 241 ENYVWKQVVDKAVTSNVASVQCDNSKFPVYKPGKCPGLGDEGTDLYDY 300

QY 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSFLFPK 330
   |||||
Db 301 EGDKVRATRVVKFPTKAHTTPWGLFYSWSTASLQSQMPSSLHPTQESTKE 360

QY 331 PKDTL-----MISRTPEVT 344
   |||||
Db 361 PNFTLHMSITFSKTPKST 379

RESULT 10

```

ABB78244  
ID ABB78244 standard; Protein; 290 AA.  
XX AC ABB78244;  
XX DT 25-NOV-2002 (first entry)  
XX

XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb290/2V.  
XX  
XX Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX

OS Unidentified.  
XX

XX WO200263003-A2.  
XX

XX 15-AUG-2002.  
XX

XX 06-FEB-2002; 2002WO-US03549.  
XX

XX 06-FEB-2001; 2001US-266838P.  
XX

XX (GEMY ) GENETICS INST LLC.  
XX

XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX

XX WPI; 2002-657537/70.  
XX

XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Disclosure; Page 11; 45pp; English.  
XX

XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which  
CC are used to produce fusion proteins with an immunoglobulin (Ig)  
CC polypeptide. The fusion proteins inhibit the adherence of platelets to  
CC leukocytes. The fusion polypeptides are useful for treating a disorder  
CC associated with platelet activation e.g. ischaemic heart disease, acute  
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,  
CC arterial thrombosis or unstable angina. They can also be used to treat  
CC vascular conditions associated with vascular inflammation, thrombosis,  
CC and angioplasty-related restenosis.  
XX

XX Sequence 290 AA;  
SQ

Query Match 54.1%; Score 1536; DB 23; Length 290;  
Best Local Similarity 100.0%; Pred. No. 1e-106;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCDKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 76

Db 1 HPICEVSKVASHLEVNCDKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 60

QY 77 NLDRCELTKLQVDTGLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136

Db 61 NLDRCELTKLQVDTGLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196

Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180

QY 197 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCNCEILYFRRLQDNANENYVWKQVVDKAVT 256

Db 181 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCNCEILYFRRLQDNANENYVWKQVVDKAVT 240

QY 257 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDYYPPEEDTEGDKVR 306

Db 241 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDYYPPEEDTEGDKVR 290

RESULT 11  
ABB78240

ID ABB78240 standard; Protein; 301 AA.  
XX AC ABB78240;  
XX

XX DT 25-NOV-2002 (first entry)  
XX

XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb302.  
XX  
XX Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX

XX Unidentified.  
XX

XX WO200263003-A2.  
XX

XX 15-AUG-2002.  
XX

XX 06-FEB-2002; 2002WO-US03549.  
XX

XX 06-FEB-2001; 2001US-266838P.  
XX

XX (GEMY ) GENETICS INST LLC.  
XX

XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX

XX WPI; 2002-657537/70.  
XX

XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Disclosure; Page 10; 45pp; English.  
XX

XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which  
CC are used to produce fusion proteins with an immunoglobulin (Ig)  
CC polypeptide. The fusion proteins inhibit the adherence of platelets to  
CC leukocytes. The fusion polypeptides are useful for treating a disorder  
CC associated with platelet activation e.g. ischaemic heart disease, acute  
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,  
CC arterial thrombosis or unstable angina. They can also be used to treat  
CC vascular conditions associated with vascular inflammation, thrombosis,  
CC and angioplasty-related restenosis.  
XX

XX Sequence 301 AA;  
SQ

Query Match 53.8%; Score 1527; DB 23; Length 301;  
Best Local Similarity 93.0%; Pred. No. 5e-106;  
Matches 292; Conservative 2; Mismatches 6; Indels 14; Gaps 1;

QY 17 HPICEVSKVASHLEVNCDKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 76

Db 1 HPICEVSKVASHLEVNCDKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 60

QY 77 NLDRCELTKLQVDTGLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136

Db 61 NLDRCELTKLQVDTGLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196

Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180

QY 197 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCNCEILYFRRLQDNANENYVWKQVVDKAVT 256

Db 181 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCNCEILYFRRLQDNANENYVWKQVVDKAVT 240

QY 257 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDYYPPEEDTEGDKVRPHTCPPCAP 316

Db 241 SNVASQCDNSDKFPVYKYPGKGCPTLGDEGDTLDYDYYPEEDTEGDKVR ----- 290

QY 317 EALGAPSVLFEPK 330

Db 291 ----ATRTWKFPK 300

RESULT 12

ABB78243

ID ABB78243 standard; Protein; 290 AA.

XX AC ABB78243;

XX DT 25-NOV-2002 (first entry)

XX DE Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb290.

XX KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;

XX KW leukocyte; platelet activation; ischaemic heart disease;

XX KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;

XX KW arterial thrombosis; angina; vascular condition; vascular inflammation;

XX KW thrombosis; angioplasty; restenosis.

XX OS Unidentified.

XX PN MO200263003-A2.

XX PD 15-AUG-2002.

XX PF 06-FEB-2002; 2002MO-US03549.

XX PR 06-FEB-2001; 2001US-266838P.

XX PA (GEMY ) GENETICS INST LLC.

XX PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;

XX DR MPI; 2002-657537/70.

XX PT New glycoprotein 1b alpha fusion polypeptides, useful for treating a

XX PT disorder associated with platelet activation e.g. ischaemic heart

XX PT disease, stroke, venous or arterial thrombosis or atherosclerosis -

XX PS Disclosure; Page 10-11; 45pp; English.

XX CC ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which

XX CC are used to produce fusion proteins with an immunoglobulin (Ig)

XX CC polypeptide. The fusion proteins inhibit the adherence of platelets to

XX CC leukocytes. The fusion polypeptides are useful for treating a disorder

XX CC associated with platelet activation e.g. ischaemic heart disease, acute

XX CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,

XX CC arterial thrombosis or unstable angina. They can also be used to treat

XX CC vascular conditions associated with vascular inflammation, thrombosis,

XX CC and angioplasty-related restenosis.

XX SQ Sequence 290 AA;

Query Match 53.8%; Score 1526; DB 23; Length 290;

Best Local Similarity 99.3%; Pred. No. 5.7e-106;

Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 76

Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 60

QY 77 NLDRCETLKLQVDTLPVLGTLDSLNQOSLPGLGTLPALTVLDVSNFRLTSLPLGAL 136

Db 61 NLDRCETLKLQVDTLPVLGTLDSLNQOSLPGLGTLPALTVLDVSNFRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLAGLENLDTLLIQ 196

Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLAGLENLDTLLIQ 180

QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNFWLNCCELLYPRRWLQDQNAENVYVWKQVDDKAVT 256

Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNFWLNCCELLYPRRWLQDQNAENVYVWKQVDDKAVT 240

QY 257 SNVASQCDNSDKFPVYKYPGKGCPTLGDEGDTLDYDYYPEEDTEGDKVR 306

Db 241 SNVASQCDNSDKFPVYKYPGKGCPTLGDEGDTLDYDYYPEEDTEGDKVR 290

RESULT 13

AAP91368

ID AAP91368 standard; peptide; 293 AA.

XX AC AAP91368;

XX DT 25-MAR-2003 (updated)

XX DT 21-MAR-1990 (first entry)

XX DE 45 kDa amino terminal tryptic fragment of glyocalicin and derivs.

XX KW Glyocalicin; von Willebrand factor; platelet membrane glycoprotein 1b;

XX KW platelet aggregation prevention; thrombosis inhibition; antithrombotic

XX KW agent

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 18..34

FT Peptide 21..35

FT Peptide 26..40

FT Peptide 26..34

FT Peptide 141..155

FT Peptide 231..245

FT Peptide 271..285

FT Peptide 281..285

XX PN EP317278-A.

XX PD 24-MAY-1989.

XX PF 16-NOV-1988; 88EP-0310799.

XX PR 17-NOV-1987; 87US-0121454.

XX PA (SCRI ) SCRIPPS CLINIC & RES FOUND.

XX PI Zimmerman TS, Ruggeri ZM, Houghten RA, Vincete V, Mohri H;

XX DR WPI; 1989-152756/21.

XX PT Proteolytic 45 KD fragment of glyocalicin and derivs. - which inhibit

XX PT binding of von Willebrand factor to platelet membrane glyco:protein, and

XX PT used as antithrombotic agents

XX PS Claim 1; Page -; 10pp; English.

XX CC The claim is for a peptide of a 45KD amino terminal tryptic fragment of

XX CC glyocalicin selected from the sequence which inhibits binding of von

XX CC Willebrand factor to platelet membrane glycoprotein 1b and related

XX CC molecules or other cells and cell matrices. Also claimed are a sequential

XX CC subset of the above (Claim 2) and specific peptides (see FT) (Claim 3)

XX CC with the same functions and any peptide of any sequential subset of amino

XX CC acids of the sequence (Claim 4). The peptides and derivs. prevent

XX CC platelet aggregation and inhibit thrombosis.

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 293 AA;

Query Match 53.8%; Score 1526; DB 10; Length 293;

Best Local Similarity 99.3%; Pred. No. 5.7e-106;

Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 76

```

|||||
1  HPICEVSKVASHLEVNCCKRNLTPALPPDLPKOTTHLHSENLIYTFSLATIMPYTRLTQ 60
|||||
77  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNLTSPLGAL 136
|||||
61  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNLTSPLGAL 120
|||||
137  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
|||||
121  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
|||||
197  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQGVVKAVT 256
|||||
181  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQGVVKAMT 240
|||||
257  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVR 306
|||||
241  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVR 290
|||||

RESULT 14
ABB78241
ID  ABB78241 standard; Protein; 302 AA.
XX  AC
XX  ABB78241;
XX
XX  25-NOV-2002 (first entry)
XX
XX  Amino acid sequence of glycoprotein Ib-alpha polypeptide GPIb302/2A.
XX
XX  Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX  leukocyte; platelet activation; ischaemic heart disease;
XX  acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX  arterial thrombosis; angina; vascular condition; vascular inflammation;
XX  thrombosis; angioplasty; restenosis.
XX
XX  Unidentified.
XX
XX  WO200263003-A2.
XX
XX  15-AUG-2002.
XX
XX  06-FEB-2002; 2002WO-US03545.
XX
XX  06-FEB-2001; 2001US-266838P.
XX
XX  (GEMY ) GENETICS INST LLC.
XX
XX  Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX
XX  WPI; 2002-657537/70.
XX
XX  New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX  disorder associated with platelet activation e.g. ischaemic heart
XX  disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX
XX  Disclosure; Page 10; 45pp; English.
XX
XX  ABB78240-45 represent glycoprotein Ib-alpha (GPIb) polypeptides, which
XX  are used to produce fusion proteins with an immunoglobulin (Ig)
XX  polypeptide. The fusion proteins inhibit the adherence of platelets to
XX  leukocytes. The fusion polypeptides are useful for treating a disorder
XX  associated with platelet activation e.g. ischaemic heart disease, acute
XX  myocardial infarction, stroke, venous thrombosis, atherosclerosis,
XX  arterial thrombosis or unstable angina. They can also be used to treat
XX  vascular conditions associated with vascular inflammation, thrombosis,
XX  and angioplasty-related restenosis.
XX
XX  Sequence 302 AA;
XX
XX  Query March 53.6%; Score 1521.5; DB 23; Length 302;
XX  Best Local Similarity 92.7%; Pred. No. 1.3e-105;
XX  Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

```

```

QY 17 HPICEVSKVASHLEVNCCKRNLTPALPPDLPKOTTHLHSENLIYTFSLATIMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCCKRNLTPALPPDLPKOTTHLHSENLIYTFSLATIMPYTRLTQ 60
QY 77 NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNLTSPLGAL 136
DB 61 NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNLTSPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQGVVKAVT 256
DB 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQGVVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVRPHTCPPCPAP 316
DB 241 SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKV----- 289
QY 317 BALGAPSVFLFPPK 330
DB 290 --AATATVVKPPTK 301

RESULT 15
AAR89436
ID  AAR89436 standard; Protein; 610 AA.
XX  AC
XX  AAR89436;
XX  25-MAR-2003 (updated)
XX  02-SEP-1996 (first entry)
XX
XX  Mutated platelet glycoprotein-Ib-alpha GPIba protein sequence.
XX
XX  Platelet glycoprotein-Ib-alpha; GPIba; mutagenesis; point mutation;
XX  von Willebrand factor; blood disorder; platelet disorder;
XX  protein engineering; Bernard-Soulier disease.
XX
XX  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  Misc-difference 57
XX  /note= "substitution from Phe in wild-typ
XX  GPIb-alpha"
XX  Region 36..200
XX  /note= "Leu rich area"
XX  Region 200..220
XX  /note= "flanking region to Leu rich area"
XX  Region 220..310
XX  /note= "hinge region"
XX  Region 310..420
XX  /note= "Ser/Thr rich area"
XX
XX  US5492809-A.
XX
XX  20-FEB-1996.
XX
XX  09-SEP-1993; 93US-0119262.
XX
XX  15-JAN-1992; 92US-0821717.
XX  07-OCT-1991; 91US-0770968.
XX  09-SEP-1993; 93US-0119262.
XX  (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX
XX  Cunningham D, Finch CN, Lyle VA, Miller JL;
XX
XX  WPI; 1996-128585/13.
XX
XX  DNA encoding platelet glyco:protein Ib alpha mutant Phe57 -

```

PT introduced into platelets to reduce aggregation and reactivity with:  
PT von Willebrand factor, also probe for diagnosis of Bernard-Soulier  
PT disease

XX

PS Disclosure; Column 21-26; 20pp; English.

XX

CC A substitution of T for C at position 259 in the DNA sequence of  
CC GPIb-alpha leads to the replacement of Phe for Ieu at  
CC residue 57 of the mature GPIb-alpha molecule. This mutated GPIb-  
CC alpha protein is less reactive with von Willebrand factor, a  
CC a characteristic of the autosomal recessive bleeding disorder,  
CC Bernard-Soulier disease.  
CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 610 AA;

```
Query Match      53.6%; Score 1521.5; DB 17; Length 610;
Best Local Similarity 82.9%; Pred. No. 3.3e-105;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;

QY 17 HPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLSSENLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLSSENLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCBLTKLQVDGTLPLVGLTDLDSNQLQSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCBLTKLQVDGTLPLVGLTDLDSNQLQSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLLO 196
Db 121 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLLO 180

QY 197 ENSLYTIPKGGFSGSHLLPFAFLHGNPWLNCCELLYFRRWLQDNAENVYVKQVDVKAVT 256
Db 181 ENSLYTIPKGGFSGSHLLPFAFLHGNPWLNCCELLYFRRWLQDNAENVYVKQVDVKAVT 240

QY 257 SNVASVQCDSNDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEJTEGDKVR----- 306
Db 241 SNVASVQCDSNDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEJTEGDKVRATRTVVVKFPT 300

QY 307 -PHTCP-----PCPA---PEALGAPSVFLPPPK--PKDTL-----MISRTPT 341
Db 301 KAHTTGWGLFYWSWTASLDSPSSUHTQBSTKEQTTPPRWTNFTLHMESITFSKTP 360

QY 342 EVT 344
Db 361 KST 363
```

Search completed: October 14, 2003, 06:30:27  
Job time : 318.154 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:31:18 ; Search time 192.642 Seconds  
(without alignments)  
444.136 Million cell updates/sec

Title: US-10-068-426-5  
Perfect score: 2839  
Sequence: 1 MFLLLLLLLSPLPHPIC.....MHEALNHYTKSLSPGK 531

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues  
Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pap.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pap.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06 NEW PUB.pap.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07 NEW PUB.pap.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08 NEW PUB.pap.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09 NEW PUB.pap.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pap.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pap.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2839	100.0	531	15 US-10-068-426-5	Sequence 5, Appli
2	2829	99.6	531	15 US-10-068-426-4	Sequence 4, Appli
3	2823	99.4	531	15 US-10-068-426-6	Sequence 6, Appli
4	2822.5	99.4	544	15 US-10-068-426-3	Sequence 3, Appli
5	2812.5	99.1	544	15 US-10-068-426-1	Sequence 1, Appli
6	2812.5	99.1	544	15 US-10-068-426-2	Sequence 2, Appli
7	1536	54.1	230	15 US-10-068-426-11	Sequence 11, Appli
8	1526.5	53.8	302	15 US-10-068-426-7	Sequence 7, Appli
9	1526	53.6	290	15 US-10-068-426-10	Sequence 10, Appli
10	1521.5	53.6	302	15 US-10-068-426-8	Sequence 8, Appli
11	1520	53.5	290	15 US-10-068-426-12	Sequence 12, Appli
12	1514	53.3	301	15 US-10-068-426-9	Sequence 9, Appli
13	1237	43.6	313	12 US-09-825-580-4	Sequence 4, Appli
14	1237	43.6	313	15 US-10-211-786-4	Sequence 4, Appli
15	1232.5	43.4	470	15 US-10-264-634-33	Sequence 33, Appli

16	1230.5	43.3	401	9 US-09-859-361-9	Sequence 9, Appli
17	1230	43.3	388	9 US-09-784-623-16	Sequence 16, Appli
18	1230	43.3	449	16 US-10-323-268-23	Sequence 23, Appli
19	1227	43.2	467	15 US-10-264-634-29	Sequence 29, Appli
20	1226.5	43.2	397	8 US-08-779-457-47	Sequence 47, Appli
21	1226.5	43.2	399	9 US-09-859-361-7	Sequence 7, Appli
22	1225.5	43.2	382	15 US-10-207-655-307	Sequence 307, Appli
23	1225.5	43.2	580	11 US-09-972-268-36	Sequence 36, Appli
24	1223	43.1	451	9 US-09-920-171-14	Sequence 14, Appli
25	1223	43.1	451	9 US-09-920-171-16	Sequence 16, Appli
26	1223	43.1	451	9 US-09-920-171-18	Sequence 18, Appli
27	1223	43.1	451	11 US-09-925-179-65	Sequence 65, Appli
28	1223	43.1	451	11 US-09-925-179-66	Sequence 66, Appli
29	1223	43.1	451	12 US-10-113-996-14	Sequence 14, Appli
30	1223	43.1	451	12 US-10-113-996-16	Sequence 16, Appli
31	1223	43.1	451	12 US-10-113-996-18	Sequence 18, Appli
32	1223	43.1	451	12 US-10-292-869-2	Sequence 2, Appli
33	1222	43.0	451	12 US-09-792-938-2	Sequence 8, Appli
34	1222	43.0	453	9 US-09-802-077-8	Sequence 8, Appli
35	1222	43.0	453	9 US-09-802-096-8	Sequence 8, Appli
36	1222	43.0	453	11 US-09-925-179-8	Sequence 25, Appli
37	1222	43.0	467	15 US-10-264-634-25	Sequence 25, Appli
38	1222	43.0	492	15 US-10-264-634-27	Sequence 27, Appli
39	1222	43.0	533	9 US-09-792-2008-4	Sequence 4, Appli
40	1222	43.0	977	12 US-10-357-653-1	Sequence 1, Appli
41	1220.5	43.0	694	10 US-09-935-868-18	Sequence 18, Appli
42	1220.5	43.0	694	12 US-10-282-162-18	Sequence 18, Appli
43	1220.5	43.0	694	15 US-10-287-035-18	Sequence 18, Appli
44	1220.5	43.0	793	10 US-09-935-868-32	Sequence 32, Appli
45	1220.5	43.0	793	12 US-10-282-162-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1

US-10-068-426-5  
; Sequence 5, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068.426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290/2V-Ig  
US-10-068-426-5

Query Match 100.0%; Score 2839; DB 15; Length 531;  
Best Local Similarity 100.0%; Pred. No. 3.9e-219;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MFLLLLLLLSPLPHPICVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY	60
Db	1	MFLLLLLLLSPLPHPICVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY	60
Qy	61	TFSLATLMPYTRLTQLNLDRCLETQLQVDGTLPLVLTGLDLSHNLQSLPLGLQTLPALTV	120

Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLLGOTLPALT 120  
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180  
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180  
Qy 181 AGLLNGLENLDTLLQENSLTYTPKGGFSGHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240  
Db 181 AGLLNGLENLDTLLQENSLTYTPKGGFSGHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240  
Qy 241 ENVYVMKQVVDKAVTSNVASVQCDNSDKFPVYKPGKCPPTLGDGDTDLDYYPEEDT 300  
Db 241 ENVYVMKQVVDKAVTSNVASVQCDNSDKFPVYKPGKCPPTLGDGDTDLDYYPEEDT 300  
Qy 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
Db 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTI 420  
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTI 420  
Qy 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480  
Db 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480  
Qy 481 VLDSGSGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531  
Db 481 VLDSGSGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531

## RESULT 2

US-10-068-426-4  
; Sequence 4, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Plaflet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPIb290-IG  
US-10-068-426-4

Query Match 99.6%; Score 2829; DB 15; Length 531;  
Best Local Similarity 99.6%; Pred. No. 2.5e-218;  
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDITILHSENLY 60  
Db 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDITILHSENLY 60  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLLGOTLPALT 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLLGOTLPALT 120

Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180  
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180  
Qy 181 AGLLNGLENLDTLLQENSLTYTPKGGFSGHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240  
Db 181 AGLLNGLENLDTLLQENSLTYTPKGGFSGHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240  
Qy 241 ENVYVMKQVVDKAVTSNVASVQCDNSDKFPVYKPGKCPPTLGDGDTDLDYYPEEDT 300  
Db 241 ENVYVMKQVVDKAVTSNVASVQCDNSDKFPVYKPGKCPPTLGDGDTDLDYYPEEDT 300  
Qy 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
Db 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTI 420  
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTI 420  
Qy 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480  
Db 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480  
Qy 481 VLDSGSGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531  
Db 481 VLDSGSGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531

## RESULT 3

US-10-068-426-6  
; Sequence 6, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Plaflet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPIb290/1A-IG  
US-10-068-426-6

Query Match 99.4%; Score 2823; DB 15; Length 531;  
Best Local Similarity 99.4%; Pred. No. 7.5e-218;  
Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDITILHSENLY 60  
Db 1 MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDITILHSENLY 60  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLLGOTLPALT 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLLGOTLPALT 120  
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180  
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180

Qy	181	AGLLNGLENDTLLLIQENS	LYTI	PKGFTGSHLLPFAFLHGNPW	CNCEILYFRFWQDNA	240
Db	181	AGLLNGLENDTLLLIQENS	SLYTI	PKGFTGSHLLPFAFLHGNPWL	CNCEILYFRFWQDNA	240
Qy	241	ENVYVMKQVVDVKAVTS	NVS	VQCUNSDKFPVYK	PGKGCPITLGDGDTLDLYYPEEDT	300
Db	241	ENVYVMKQVVDVAMTS	NVS	VQCUNSDKFPVYK	PGKGCPITLGDGDTLDLYYPEEDT	300
Qy	301	EGDKVRPHTCPPCPAPE	ALGAPSV	FLPFPKPD	TLMSIRPTEVTCVVVDVSHEDPEVKEN	360
Db	301	EGDKVRPHTCPPCPAPE	ALGAPSV	FLPFPKPD	TLMSIRPTEVTCVVVDVSHEDPEVKEN	360
Qy	361	WYVDGVEVHNATKPRE	QVNSTYRW	SVLTVLHQDWLNGKE	YCKCKVSNKALPVIPIEKT	420
Db	361	WYVDGVEVHNATKPRE	QVNSTYRW	SVLTVLHQDWLNGKE	YCKCKVSNKALPVIPIEKT	420
Qy	421	SKAGQPREPQVYTLPP	SREEMTKNQVSLTCL	VKGFP	PSDIAVEWESNGQPENNYKTTTP	480
Db	421	SKAGQPREPQVYTLPP	SREEMTKNQVSLTCL	VKGFP	PSDIAVEWESNGQPENNYKTTTP	480
Qy	481	VLDSDGSFFLYSKLTV	DKSRWQQGVN	FSCSVMHEALNNHY	TQKSLSLSPGK	531
Db	481	VLDSDGSFFLYSKLTV	DKSRWQQGVN	FSCSVMHEALNNHY	TQKSLSLSPGK	531

```

RESULT 4
US-10-068-426-3
; Sequence 3, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumař, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Plati let Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPIb302/4X-Ig
US-10-068-426-3

```

```

181 181  AGLLNGLENLDTLLLRNSLTYTPKGFSGSHLLPFAFLHGNPMLCNCELLYFRFRWLDNA 2410
Qy 241  ENVYVKQVVDVKATSNVASVQCDNSDKFPVYKPGKGCPTLGDGSDTLDLYDYYPEEDT 300
Db 241  ENVYVKQVVDVKATSNVASVQCDNSDKFPVYKPGKGCPTLGDGSDTLDLYDYYPEEDT 300
Qy 301  EGDKV-----RPHTCPPCPAPEALGAPSVFLFPKPKOTLMISRTPEVTCV 347
Db 301  EGDKVAATATVVKFPTKARPHTCPPCPAPEALGAPSVFLFPKPKOTLMISRTPEVTCV 360
Qy 348  VDVSHEDEPVKFNWVDCGVEVHNAKTPRBEQYNSTYRVVSVLTVLQDMLNGKEYKCKV 407
Db 361  VDVSHEDEPVKFNWVDCGVEVHNAKTPRBEQYNSTYRVVSVLTVLQDMLNGKEYKCKV 420
Qy 408  SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
Db 421  SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
Qy 468  NQGPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQOGNVFSCSMGHEALHNHYTQKSLSL 527
Db 481  NQGPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQOGNVFSCSMGHEALHNHYTQKSLSL 540
Qy 528  SPGK 531
Db 541  SPGK 544

RESULT 5
US-10-068-426-1
; Sequence 1, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: Gplb302-Ig
US-10-068-426-1

```

Query Match	99.4%	Score 2822.5	DB 15	Length 544
Best Local Similarity	97.6%	Pred. No. 8.5e-218		
Matches 531	Conservative 0	Mismatches 0	Indels 13	Gaps 1
QY	1	MPILLULLLLPSPLPHPICEVSKVASHLEVNC DKRNLTALPPDLPKDPTTILHLSENLLY	60	
Db	1	MPILLULLLLPSPLPHPICEVSKVASHLEVNC DKRNLTALPPDLPKDPTTILHLSENLLY	60	
QY	61	TSLSATLMFYTRLTQLNLDRCBELTKLOVDGTLPVLTGLDLSHNOLQSLPLGGTLPALT	120	
Db	61	TSLSATLMFYTRLTQLNLDRCBELTKLOVDGTLPVLTGLDLSHNOLQSLPLGGTLPALT	120	
QY	121	LOVSFNRLTSLPLGALRGELQELLYKGNELKTLPPGLLTPTTPKLEKSLANNNTL	180	
Db	121	LOVSFNRLTSLPLGALRGELQELLYKGNELKTLPPGLLTPTTPKLEKSLANNNTL	180	
QY	181	AGLLNGLENLDTLLIQNSLVYTPKGFSGSHLLPFAFHGNPWLNCNCEILLYFRRLQDNA	240	

Query Match	99.1%	Score 2812.5;	DB 15;	Length 544;
Best Local Similarity	97.2%	Pred. No. 5.4e-217;		
Matches 529;	Conservative 1;	Mismatches 1;	Indels 13;	Gaps 1
QY	1	MPLLLLLLLLPSLPHPHPICEVSKVASHLEVNC	KRNLTALPDLPKD	TTILHSENLLY 60
Db	1	MPLLLLLLLLPSLPHPHPICEVSKVASHLEVNC	KRNLTALPDLPKD	TTILHSENLLY 60
QY	61	TSLSLATMPYTRLTQNLNDRCELTKLOVDG	TLPVLTGLDLSHNQLOSLP	LGGTLPALTY 120
Db	61	TSLSLATMPYTRLTQNLNDRCELTKLOVDG	TLPVLTGLDLSHNQLOSLP	LGGTLPALTY 120
QY	121	LDVSNFNRLTSLPGLARGLGELQELYLKGNEL	KLTLPGLLTPTPKLEKLSLANNLT	TELP 180
Db	121	LDVSNFNRLTSLPGLARGLGELQELYLKGNEL	KLTLPGLLTPTPKLEKLSLANNLT	TELP 180
QY	181	AGLLNGLENLDTLLAORNSLYTIPKGFSGSHLP	PFAPFLHGNPMLCNCE	TIIPRRPWODNA 240



Db 181 AGLANGLENLDTLQLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Qy 241 ENVYVWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYVWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKV-----RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347  
Db 301 EGDKVRAITVVVFTKARPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
Qy 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGYKCKV 407  
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGYKCKV 420  
Qy 408 SNKALPVPTEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEMES 467  
Db 421 SNKALPVPTEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEMES 480  
Qy 468 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHEALHNHYTKQSL 527  
Db 481 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHEALHNHYTKQSL 540  
Qy 528 SPKG 531  
Db 541 SPKG 544

RESULT 6  
US-10-068-426-2  
; Sequence 2, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: GPIb302/2A-Ig  
US-10-068-426-2

Query Match 99.1%; Score 2812.5; DB 15; Length 544;  
Best Local Similarity 97.2%; Pred. No. 5.4e-217;  
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;  
Qy 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60  
Db 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60  
Qy 61 TFSLATLAPYRTELTLQNLDRCELTKLQVDGTLPLVLTGLDLSHNQLSQSLPLGQTLPALTY 120  
Db 61 TFSLATLAPYRTELTLQNLDRCELTKLQVDGTLPLVLTGLDLSHNQLSQSLPLGQTLPALTY 120  
Qy 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGILLTPPKLEKLSANNLTPLP 180  
Db 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGILLTPPKLEKLSANNLTPLP 180

Qy 181 AGLANGLENLDTLQLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Db 181 AGLANGLENLDTLQLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Qy 241 ENVYVWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYVWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKV-----RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347  
Db 301 EGDKVAATATVVVFTKARPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
Qy 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGYKCKV 407  
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGYKCKV 420  
Qy 408 SNKALPVPTEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEMES 467  
Db 421 SNKALPVPTEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEMES 480  
Qy 468 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHEALHNHYTKQSL 527  
Db 481 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHEALHNHYTKQSL 540  
Qy 528 SPKG 531  
Db 541 SPKG 544

RESULT 7  
US-10-068-426-11  
; Sequence 11, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 290  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(290)  
; OTHER INFORMATION: GPIb290/2V  
US-10-068-426-11

Query Match 54.1%; Score 1536; DB 15; Length 290;  
Best Local Similarity 100.0%; Pred. No. 4.8e-115;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTFSLATLMPYRRLTQL 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLYTFSLATLMPYRRLTQL 60  
Qy 77 NDRCSLTKLQVDGTLPLVLTGLDLSHNQLSQSLPLGQTLPALTYLDVSNRNLTSPLGAL 136  
Db 61 NDRCSLTKLQVDGTLPLVLTGLDLSHNQLSQSLPLGQTLPALTYLDVSNRNLTSPLGAL 120  
Qy 137 RGELQELYLKGNELKTLPPGILLTPPKLEKLSANNLTPLPAGLNGLENLDTLLIQ 196  
Db 121 RGELQELYLKGNELKTLPPGILLTPPKLEKLSANNLTPLPAGLNGLENLDTLLIQ 180

Qy	197	ENSLYTI	PKGFGSGHLLPFAFLHGNPMWLCNCELLYFRRWLODNNAENTVWKQVVDVKAVT	256
Db	181	ENSLYTI	PKGFGSGHLLPFAFLHGNPMWLCNCELLYFRRWLODNNAENYVWKQVVDVKAVT	240
Qy	257	SNVASQ	CNSDKFPYKYIPGKGCTPLGDEGDTLDYYPSEEDTEGDKVR	306
Db	241	SNVASQ	CNSDKFPYKYIPGKGCTPLGDEGDTLDYYPSEEDTEGDKVR	290

## RESULT 8

```

US-10-068-426-7
; Sequence 7, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GPLb302
US-10-068-426-7

```

RESIT.T 9

US-10-068-426-10  
; Sequence 10, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:

```

; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

Query Match      53.6%; Score 1521.5; DB 15; Length 302;
Best Local Similarity 92.7%; Pred. No. 7.4e-114;
Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 256
DB 181 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKVR 306
DB 241 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKVR 290

RESULT 12
US-10-068-426-9
; Sequence 9, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068.426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(301)
; OTHER INFORMATION: GP1b/4X
US-10-068-426-9

Query Match      53.3%; Score 1514; DB 15; Length 301;
Best Local Similarity 93.0%; Pred. No. 2.9e-113;
Matches 292; Conservative 1; Mismatches 7; Indels 14; Gaps 2;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 256
DB 181 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKVR 316
DB 241 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKV----- 288
QY 317 EALGAPSVFLFPK 330
DB 289 --AATATVVKFPTK 300

; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

Query Match      53.6%; Score 1521.5; DB 15; Length 302;
Best Local Similarity 92.7%; Pred. No. 7.4e-114;
Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 256
DB 181 ENSLYTIPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNANVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKVR 316
DB 241 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDYDYPEEDTEGDKV----- 289
QY 317 EALGAPSVFLFPK 330
DB 290 --AATATVVKFPTK 301

RESULT 11
US-10-068-426-12
; Sequence 12, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068.426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: GB1b290/1A
US-10-068-426-12

Query Match      53.5%; Score 1520; DB 15; Length 290;
Best Local Similarity 99.0%; Pred. No. 9.2e-114;
Matches 287; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLLGOTLPALTVDVSFNRLTSLPLGAL 136

```

RESULT 13

US-09-825-580-4  
; Sequence 4, Application US/09825580  
; Publication No. US20030166521A1  
; GENERAL INFORMATION:  
; APPLICANT: Eppihimer, Michael J.  
; APPLICANT: Schaub, Robert G.  
; APPLICANT: Harris, Alan  
; TITLE OF INVENTION: Inhibition of Thrombosis by Treatment with  
; FILE REFERENCE: GFN-5398  
; CURRENT APPLICATION NUMBER: US/09/825,580  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/193,787  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-825-580-4

Query Match 43.6%; Score 1237; DB 12; Length 313;  
Best Local Similarity 97.0%; Pred. No. 5e-91;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 296 PEEDTEGDKVRHTCCPCPAPEALGAPSVLFPKPKDTLMISRTPVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARHTCCPCPAPEALGAPSVLFPKPKDTLMISRTPVTCVVVDVSHEDP 137  
  
QY 356 EVKFNMYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 197  
  
QY 416 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 475  
DB 198 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 257  
  
QY 476 KTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSLSPGK 531  
DB 258 KTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSLSPGK 313

RESULT 14

US-10-211-786-4  
; Sequence 4, Application US/10211786  
; Publication No. US20030083256A1  
; GENERAL INFORMATION:  
; APPLICANT: Michael J. Eppihimer  
; APPLICANT: Robert G. Schaub  
; APPLICANT: Ronald Tuma  
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM  
; FILE REFERENCE: 8702.0099-00000  
; CURRENT APPLICATION NUMBER: US/10/211,786  
; PRIOR FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: 60/309,816  
; PRIOR FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-211-786-4

Query Match 43.6%; Score 1237; DB 15; Length 313;  
Best Local Similarity 97.0%; Pred. No. 5e-91;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 296 PEEDTEGDKVRHTCCPCPAPEALGAPSVLFPKPKDTLMISRTPVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARHTCCPCPAPEALGAPSVLFPKPKDTLMISRTPVTCVVVDVSHEDP 137

QY 356 EVKFNMYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 197  
  
QY 416 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 475  
DB 198 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 257  
  
QY 476 KTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSLSPGK 531  
DB 258 KTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSLSPGK 313

RESULT 15

US-10-264-634-33  
; Sequence 33, Application US/10264634  
; Publication No. US20030108549A1  
; GENERAL INFORMATION:  
; APPLICANT: Donaldson, Debra et al.  
; TITLE OF INVENTION: Methods and Compositions for Modulating Interleukin-21 Receptor  
; FILE REFERENCE: GI5320-P3  
; CURRENT APPLICATION NUMBER: US/10/264,634  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: 09/040,005  
; PRIOR FILING DATE: 1998-03-17  
; PRIOR APPLICATION NUMBER: 09/560,766  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/569,384  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: 09/972,218  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/373,746  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 33  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Human  
US-10-264-634-33

Query Match 43.4%; Score 1232.5; DB 15; Length 470;  
Best Local Similarity 74.6%; Pred. No. 2e-90;  
Matches 241; Conservative 16; Mismatches 39; Indels 27; Gaps 3;

QY 220 GNPWLCNCEILYFRWLQDNAENVYVWKQVVDKAVTSNVASVQCDNSDKFPVYKPGK 279  
DB 164 GDPMAVSP-----RRKLIS-----VDSRSVSLLEFRKQSSYELQVRAGMPG 207  
  
QY 280 C-----PTLGDGEGDTLDVYYPEDTEGDKVRPHTCCPCPAPEALGAPSVLFP 328  
DB 208 SSYQGTWSESDPVIFFQOSEELKEGWNDDDDKSGDKTHTCCPCPAPEALGAPSVLFP 267  
  
QY 329 PKPKDTLMISRTPVTCVVVDVSHEDPEVKFNMYVDGVEVHNKTKPREEQYNSTYRVVS 388  
DB 268 PKPKDTLMISRTPVTCVVVDVSHEDPEVKFNMYVDGVEVHNKTKPREEQYNSTYRVVS 327  
  
QY 389 VLTVLHQDWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPQVYTLPPSRREMTKNQVS 448  
DB 328 VLTVLHQDWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPQVYTLPPSRREMTKNQVS 387  
  
QY 449 LTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFS 508  
DB 388 LTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFS 447

QY 509 CSMVHEALHNYTKSLSLSPGK 531  
DB 448 CSMVHEALHNYTKSLSLSPGK 470

Search completed: October 14, 2003, 06:48:24  
Job time : 195.642 secs

-----

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: October 14, 2003, 06:17:26 ; Search time 21.734 Seconds  
(without alignments)  
1033.731 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPLLILLLLPLPHPIC.....MHEALHNYTOKLSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first: 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528.5	53.8	320	1	US-07-613-083B-1
2	1521.5	53.6	610	1	US-07-821-717B-6
3	1521.5	53.6	610	1	US-08-119-262B-6
4	1521.5	53.6	610	1	US-08-135-929A-11
5	1521.5	53.6	610	1	US-08-234-265A-11
6	1237	43.6	313	3	US-08-713-556F-36
7	1230	43.3	388	3	US-09-131-247-16
8	1230	43.3	449	3	US-08-897-236-23
9	1230	43.3	449	3	US-09-500-253B-23
10	1223	43.1	451	2	US-08-887-352B-14
11	1223	43.1	451	2	US-08-887-352B-16
12	1223	43.1	451	2	US-08-887-352B-18
13	1223	43.1	451	3	US-08-466-151-65
14	1223	43.1	451	3	US-09-109-207C-14
15	1223	43.1	451	3	US-09-109-207C-16
16	1223	43.1	451	3	US-09-109-207C-18
17	1223	43.1	451	3	US-09-282-505-2
18	1223	43.1	451	3	US-09-054-255-2
19	1223	43.1	451	3	US-09-296-005-14
20	1223	43.1	451	3	US-09-296-005-16
21	1223	43.1	451	3	US-09-296-005-18
22	1223	43.1	451	4	US-09-282-846-2
23	1223	43.1	451	4	US-09-680-145-2
24	1222	43.0	453	3	US-08-466-151-8
25	1222	43.0	453	4	US-08-466-163B-8
26	1222	43.0	977	4	US-09-590-656-1
27	1222	43.0	977	4	US-09-733-764-1

Sequence 18, Appl  
Sequence 32, Appl  
Sequence 30, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 15, Appl  
Sequence 9, Appl  
Sequence 26, Appl  
Sequence 24, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-07-613-083B-1  
; Sequence 1, Application US/07613083B  
; Patent No. 5340727  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Ware, Jerry, inventors  
; APPLICANT: on behalf of Scripps Clinic and Research  
; APPLICANT: Foundation  
; TITLE OF INVENTION: GP1b' Fragments and Recombinant  
; TITLE OF INVENTION: DNA Expression Vectors  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scripps Clinic and Research  
; ADDRESSEE: Foundation  
; STREET: 10666 No. 5340727th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb  
; COMPUTER: AST Bravo IBM PC comp. (386SX)  
; OPERATING SYSTEM: MS DOS version 3.2  
; SOFTWARE: Wordperfect 5.1 conv. to ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/613,083B  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: This appl. is a c-i-p of  
; APPLICATION NUMBER: U.S. 07/470,674  
; FILING DATE: 04-Jan-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barron, Alexis  
; REGISTRATION NUMBER: 22,702  
; REFERENCE/DOCKET NUMBER: P16,569-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 923-4466  
; TELEFAX: (215) 923-2189  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320  
; TYPE: AMINO ACID  
; STRANDEDNESS: No. 5340727 applicable  
; TOPOLOGY: Linear  
US-07-613-083B-1

Query Match 53.8%; Score 1528.5; DB 1; Length 320;  
Best Local Similarity 95.1%; Pred. No. 2.3e-129;  
Matches 291; Conservative 1; Mismatches 3; Indels 11; Gaps 1;

17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76  
1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60  
77 NLDRCETLKQVDTLPVLGTLDLSHNQLQSLPLGQTLPALTVLDVSNRNLTSPLGAL 136  
61 NLDRCETLKQVDTLPVLGTLDLSHNQLQSLPLGQTLPALTVLDVSNRNLTSPLGAL 120  
137 RGLGELQELYLKGNELKTLPPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 196  
121 RGLGELQELYLKGNELKTLPPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 180  
197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVVDVKAVT 256  
181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVVDVKAVT 240  
257 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYYPEDTEGDKVR----- 306  
241 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYYPEDTEGDKVRATRTVVVKFPT 300  
307 -PHTCP 311  
301 KAHTTP 306

## RESULT 2

US-07-821-717B-6  
; Sequence 6, Application US/07821717B  
; Patent No. 5298239  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/821,717B  
; FILING DATE: 15-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.

AUTHORS: Papayannopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human  
TITLE: platelet glycoprotein Ib: A transmembrane protein with homology  
TITLE: to leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic  
TITLE: peptides that block the binding of von Willebrand factor to the  
TITLE: platelet membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-07-821-717B-6

Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76  
Db 1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCETLKQVDTLPVLGTLDLSHNQLQSLPLGQTLPALTVLDVSNRNLTSPLGAL 136  
Db 61 NLDRCETLKQVDTLPVLGTLDLSHNQLQSLPLGQTLPALTVLDVSNRNLTSPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 196  
Db 121 RGLGELQELYLKGNELKTLPPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVVDVKAVT 256  
Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVVDVKAVT 240  
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYYPEDTEGDKVR----- 306  
Db 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYYPEDTEGDKVRATRTVVVKFPT 300  
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPPK---PKDTL-----MISRT 341  
Db 301 KAHTTPWGLFYSWSTASLDSOMPSSLHPTQESTKEQTFFPRWTFNFTLHMEITFSKTP 360  
QY 342 EVT 344  
Db 361 KST 363

## RESULT 3

US-08-119-262B-6  
; Sequence 6, Application US/08119262B  
; Patent No. 5492809  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051

CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,262B  
FILING DATE: 09-SEP-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/821,717  
FILING DATE: 15-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/22  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Lopez, Jose A.  
AUTHORS: Chung, Dominic W.  
AUTHORS: Fujikawa, Kazuo  
AUTHORS: Hagen, Frederick S.  
AUTHORS: Papayanopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human platelet glycoprotein Ib: A transmembrane protein  
TITLE: leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic peptides that block the binding of von Willebrand factor to the platelet glycoprotein Ib  
TITLE: membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-08-119-262B-6  
Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKOTTLHLSNLLYTFSLATLMPYTRLTOL 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKOTTLHLSNLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDTLPLVGLTDLSHNQLQSLPLGQTLPALTVLDVSNFRLTSLPLGAL 136  
Db 61 NLDRCCLTKLQVDTLPLVGLTDLSHNQLQSLPLGQTLPALTVLDVSNFRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPVGLINGLENLDTLLQ 196

Db 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPVGLINGLENLDTLLQ 180  
QY 197 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDNAENVYVWKQVVDYKAVT 256  
Db 181 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDNAENVYVWKQVVDYKAVT 240  
QY 257 SNVASVQCDNSDKFPVYKYGKGCPTLGDGDTLDLYDYDEEDTEGDKVR----- 306  
Db 241 SNVASVQCDNSDKFPVYKYGKGCPTLGDGDTLDLYDYDEEDTEGDKVRATRVVVKFT 300  
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPPK--PKDTL-----MISRTPT 341  
Db 301 KAHTTPMGLFYWSASTASLDSPSSSLHPTQESTKEQTTPFPRTNPTFLHMSITFSKTP 360  
QY 342 EVT 344  
Db 361 KST 363  
RESULT 4  
US-08-135-929A-11  
Sequence 11, Application US/08135929A  
Patent No. 5593959  
GENERAL INFORMATION:  
APPLICANT: Miller, Jonathan L.  
APPLICANT: Cunningham, David  
APPLICANT: Lyle, Vicki A.  
APPLICANT: Finch, Clara N.  
APPLICANT: Pincus, Matthew R.  
TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha  
TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/135,929A  
FILING DATE: 14-OCT-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/23  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
TELEX: 978450  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-135-929A-11  
Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKOTTLHLSNLLYTFSLATLMPYTRLTOL 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKOTTLHLSNLLYTFSLATLMPYTRLTOL 60





APPLICATION NUMBER: US 08/316,305  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/428,734  
FILING DATE: 25-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWN, SCOTT A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI 5213F  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-713-556F-36

Query Match 43.6%; Score 1237; DB 3; Length 313;  
Best Local Similarity 97.0%; Pred. No. 3.8e-103;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 296 PBDTEGDKVRPHPTCPPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARPHPTCPPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 137  
QY 356 EVKFNMYVDGVEVHNNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVP 197  
QY 416 IETTSKAKGQPEPOVYTLPPSRESMTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNY 475  
DB 198 IETTSKAKGQPEPOVYTLPPSRESMTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNY 257  
QY 476 KITPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSQVWMEALHNHYTKQSLSPGK 531  
DB 258 KITPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSQVWMEALHNHYTKQSLSPGK 313

RESULT 7  
US-09-131-247-16  
Sequence 16, Application US/09131247  
Patent No. 6294170  
GENERAL INFORMATION:  
APPLICANT: Boone, Thomas C.  
APPLICANT: Herzhenson, Susan  
APPLICANT: Bevilacqua, Michael P.  
APPLICANT: Collins, David S.  
TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
DISEASES  
FILE REFERENCE: A-365F  
CURRENT APPLICATION NUMBER: US/09/131,247  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: 60/055,185  
EARLIER FILING DATE: 1997-08-08  
EARLIER APPLICATION NUMBER: PCT/US 97/02131  
EARLIER FILING DATE: 1997-02-10  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Human  
US-09-131-247-16  
Query Match 43.3%; Score 1230; DB 3; Length 388;  
Best Local Similarity 64.5%; Pred. No. 2.3e-102;  
Matches 249; Conservative 20; Mismatches 69; Indels 48; Gaps 6;  
QY 171 LANNLTLPAGLNG-----LENLDTLLQENSLYTIPIKPGFSGSHLLPFAFLHGNPMLC 225

DB 26 LRNN---QLVAGYLQGNVNLLEEKIDVVPVIEPHAL-----FLGIHGGKMC 68  
QY 226 NCEILYFRRLQDNNAENYVWKQVVDKAVTSNVASVQCDNSDKFPVYKPGKCP---- 281  
DB 69 SCVKSGETRLQLEAVN-----ITDISENKKQDKRFAFIRSDSGPTTSFESAACPGWFL 122  
QY 282 -----TLGDEGDTLDYYPPEEDTEGD---KVRPHTCPPCPAPEALGAPSVF 325  
DB 123 CTAMEADQPVSLTNMPDEGVMVTKFYQDEDAAEAPKSSDKTHTCTCPAPABELLGGPSV 182  
QY 326 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYR 385  
DB 183 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYR 242  
QY 386 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVTEKTSKAKGQPEPOVYTLPPSREEMTKN 445  
DB 243 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVTEKTSKAKGQPEPOVYTLPPSREEMTKN 302  
QY 446 QVSLTCLVKGFYPSDIAVEMESNGQPENNYTKTTPVLDSDGSFFLYSLKLTVDKSRWQQGN 505  
DB 303 QVSLTCLVKGFYPSDIAVEMESNGQPENNYTKTTPVLDSDGSFFLYSLKLTVDKSRWQQGN 362  
QY 506 VFSCSVMEALHNHYTKQSLSPGK 531  
DB 363 VFSCSVMEALHNHYTKQSLSPGK 388  
RESULT 8  
US-08-897-236-23  
Sequence 23, Application US/08897236A  
Patent No. 6075007  
GENERAL INFORMATION:  
APPLICANT: Regeneron Pharmaceuticals, Inc.  
TITLE OF INVENTION: Modified Dorsal Tissue Affecting Factor and Composition  
FILE REFERENCE: REG 133  
CURRENT APPLICATION NUMBER: US/08/897,236A  
CURRENT FILING DATE: 1997-07-17  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 23  
LENGTH: 449  
TYPE: PRT  
ORGANISM: human  
US-08-897-236-23  
Query Match 43.3%; Score 1230; DB 3; Length 449;  
Best Local Similarity 56.1%; Pred. No. 2.8e-102;  
Matches 273; Conservative 21; Mismatches 95; Indels 98; Gaps 13;  
QY 105 LQSLPPLGQTLPALTVLDVSNRLTSLPLGALRGU----- 139  
DB 1 MERCPSLGVTLIALVWV---LGLRATPAGQGHYLRPAPSDNLPVLVDLIEHPDIPDP 56  
QY 140 --GELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLNLTLLIQE 197  
DB 57 KEKDLNETLRLSLGCHYDPCFMATSPEDPFGGGG-----AAGGAEDLAEQLQLRQR 111  
QY 198 NSLYTIPIKPGFSGHLLPFAFLHG-----NPWLCN---CEILYFRRLQDNNAENYVWKQV 249  
DB 112 -----PSGAMPSEIKGLEFSEGLAQGLQWLWSQTFPCVLY--AW---NDLSRFWPRY 160  
QY 250 VDV-----KAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDT 289  
DB 161 VKGSCFSKRSVPEGMVCKPSKSVHLTVLRWCQ-----RRGQRCGMIPQ--- 209  
QY 290 DLYDYTP-----BEDTEGDKVRPHPTCPPEALGAPSVLPFPKPKDTLMISRTPEVT 344  
DB 210 -----YPIISECKSCSGDKT--HTCPPEALGAPSVLPFPKPKDTLMISRTPEVT 262  
QY 345 CVVDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYK 404  
DB 263 CVVDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYK 322

QY 405 CKVSNKALPVPKIEKTSKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 464  
Db 323 CKVSNKALPAPIKTIISKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 382  
QY 465 WESNGQPNKYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFSVCSVMHEALHNNHYTKQS 524  
Db 383 WESNGQPNKYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFSVCSVMHEALHNNHYTKQS 442  
QY 525 LSLSPGK 531  
Db 443 LSLSPGK 449

RESULT 9  
US-09-500-253B-23  
; Sequence 23, Application US/09500253B  
; Patent No. 6500640  
; GENERAL INFORMATION:  
; APPLICANT: Regeneron Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Modified Dorsal Tissue Affecting Factor and Composition  
; FILE REFERENCE: REG 133-2  
; CURRENT APPLICATION NUMBER: US/09/500.253B  
; FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Human  
US-09-500-253B-23

Query Match 43.3%; Score 1230; DB 4; Length 449;  
Best Local Similarity 56.1%; Pred. No. 2.8e-102;  
Matches 273; Conservative 21; Mismatches 95; Indels 98; Gaps 13;  
QY 105 LQSLPLLGQTLPALTVLDVSNFRLTSLPLGALRGL----- 139  
Db 1 MERCPSLGVTLYALVVV----LGLRATPAGGOVLIHTRAPSDNLPVLDLIEHPDIFDP 56  
QY 140 --GELQELYKGNELKTLPPGLLTPTPKLEKLSANNLTLTLPAGLNGLENLDTLLQ 197  
Db 57 KEXDLNETLLRSLGCHYDPGMATSPEDRPGGGG-----AAGGAEDLAELDQLLR 111  
QY 198 NSLYTIKPGFGSHLLPFAFLHG-----NPLCN---CEILYFRRWLQDNAENVYWKQV 249  
Db 112 -----PSGAMPSEIKGLEFSEGLAQGLQMLWMSQTFCPLY--AW---NDLGSRFNRY 160  
QY 250 VDV-----KAVTSNVALSVQCDNSDKFPVYKYPKGCGPTLGDEGT 289  
Db 161 VKVGSFCFSKSCSVPEGMVCKPSKVHLTVLRWRCQ-----RRGGRQCGWPIQ--- 209  
QY 290 DLYDYYP-----EDTSGDKVVRHTCPCPAPEALGAPSVFLPPPKKDTLMISRTPEVT 344  
Db 210 -----YPII1SECKSCSGDKT--HTCPCPAPEALGGSVFLPPPKKDTLMISRTPEVT 262  
QY 345 CVVVDVSHEDPEVKFNWVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKEYK 404  
Db 263 CVVVDVSHEDPEVKFNWVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKEYK 382  
QY 405 CKVSNKALPVPKIEKTSKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 464  
Db 323 CKVSNKALPAPIKTIISKAGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 382  
QY 465 WESNGQPNKYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFSVCSVMHEALHNNHYTKQS 524  
Db 383 WESNGQPNKYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFSVCSVMHEALHNNHYTKQS 442  
QY 525 LSLSPGK 531  
Db 443 LSLSPGK 449

RESULT 11

RESULT 10  
US-08-887-352B-14  
; Sequence 14, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
; TITLE OF INVENTION: Improved Anti-IGE Antibodies and Method of  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887.352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-14

Query Match 43.1%; Score 1223; DB 2; Length 451;  
Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
QY 188 ENLDLTLQENSLYTIKGF-----GSHLLPFAFLHGNPMLCNCILYFRRWLQDNAENV 243  
Db 74 DSKNTFYLOMNSLRADRTAVYTCARGSHY----FGH-----HFAVMGQGTTLTV 119  
QY 244 -----YWKQVVDVVKAVTSNVALSVQCDNSDKFP-----VYKYPGKGC 280  
Db 120 SSASTKGPSVFLPAPSSKSTSGTAALGCLVKDYFPFVTVSWNSGALTSGVHTFPA--- 176  
QY 281 PTLGDEG-----DTOLY-----DYPEBDEGDKVTP-----HTCPCPA 315  
Db 177 -VLQSSGLYSLSSVTVFSSSLGTQYICNVNHPKSNTKVDKCKVPSKCDKTHTCPCPA 235  
QY 316 PEALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 375  
Db 236 PELLGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 295  
QY 376 REEQYNSTYRVVSVLTVLHODWLNKEYCKCKYKSNKALPVPKIEKTSKAGQPREQVYTL 435  
Db 296 REEQYNSTYRVVSVLTVLHODWLNKEYCKCKYKSNKALPAPIKTIISKAGQPREQVYTL 355  
QY 436 PPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPNKYKTPPVLDSDGSGFFLYSKLT 495  
Db 356 PPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPNKYKTPPVLDSDGSGFFLYSKLT 415  
QY 496 VDKSRWQGNVFSVCSVMHEALHNNHYTKQSLSPGK 531  
Db 416 VDKSRWQGNVFSVCSVMHEALHNNHYTKQSLSPGK 451



Patent No. 6037453  
 GENERAL INFORMATION:  
 APPLICANT: Jardieu, Paula M.  
 TITLE OF INVENTION: Immunoglobulin Variants  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatIn (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/466,151  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/466163  
 FILING DATE: 06-Jun-1995  
 APPLICATION NUMBER: 08/405617  
 FILING DATE: 15-MAR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/185899  
 FILING DATE: 26-JAN-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/879495  
 FILING DATE: 07-MAY-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/744768  
 FILING DATE: 14-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svoboda, Craig G.  
 REGISTRATION NUMBER: 39,044  
 REFERENCE/DOCKET NUMBER: P0718P2C1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-1489  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 65:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 451 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-466-151-65

Query Match 43.1%; Score 1223; DB 3; Length 451;  
 Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
 Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
 QY 188 ENLDTLLQNSLYTIPKGF---GSHLLPFAFLHGNPWLNCNCEILYFRRLQDANV 243  
 DB 74 DSKNTFYLQNSLRAEDTAVYICARGSHY----FGH---W-----HFAVWGQGLTVTV 119  
 QY 244 -----YVMQVVDVKAFTSNVASQCDNSDKFP-----VYKYPGKGC 280  
 DB 120 SSASTKGPSVFPLAPSSKSTSGTAAALCLVKDYFPEPTVTSWNSGALTSGVHTFPA--- 176  
 QY 281 PTLGDEG-----DTDLY-----DYYPEDETEGDKVRP-----HTCPPCPA 315  
 DB 177 -VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKVEPKSCDKTHCTCPCPA 235  
 QY 316 PEALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 375  
 DB 236 PELLGGPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 295  
 QY 376 REQYNSTRYVSVLTVLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 435  
 DB 296 REEQYNSTRYVSVLTVLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 355  
 QY 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 495  
 DB 356 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 451

RESULT 15  
 US-09-109-207C-16  
 Sequence 16, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept

QY 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 495  
 DB 356 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 451  
 RESULT 14  
 US-09-109-207C-14  
 Sequence 14, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept  
 FILE REFERENCE: P1123R1  
 CURRENT APPLICATION NUMBER: US/09/109,207C  
 CURRENT FILING DATE: 1998-06-30  
 PRIOR APPLICATION NUMBER: US 60/051,554  
 PRIOR FILING DATE: 1997-07-03  
 NUMBER OF SEQ ID NOS: 44  
 SEQ ID NO 14  
 LENGTH: 451  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 NAME/KEY: Artificial  
 LOCATION: 1-451  
 OTHER INFORMATION: Heavy chain sequence derived from MAE11  
 US-09-109-207C-14

Query Match 43.1%; Score 1223; DB 3; Length 451;  
 Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
 Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
 QY 188 ENLDTLLQNSLYTIPKGF---GSHLLPFAFLHGNPWLNCNCEILYFRRLQDANV 243  
 DB 74 DSKNTFYLQNSLRAEDTAVYICARGSHY----FGH---W-----HFAVWGQGLTVTV 119  
 QY 244 -----YVMQVVDVKAFTSNVASQCDNSDKFP-----VYKYPGKGC 280  
 DB 120 SSASTKGPSVFPLAPSSKSTSGTAAALCLVKDYFPEPTVTSWNSGALTSGVHTFPA--- 176  
 QY 281 PTLGDEG-----DTDLY-----DYYPEDETEGDKVRP-----HTCPPCPA 315  
 DB 177 -VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKVEPKSCDKTHCTCPCPA 235  
 QY 316 PEALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 375  
 DB 236 PELLGGPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 295  
 QY 376 REQYNSTRYVSVLTVLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 435  
 DB 296 REQYNSTRYVSVLTVLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 355  
 QY 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 495  
 DB 356 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSGSEFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNNHYTKQSLSPGK 451

RESULT 15  
 US-09-109-207C-16  
 Sequence 16, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept

```
FILE REFERENCE: P1123R1
CURRENT APPLICATION NUMBER: US/09/109,207C
CURRENT FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/051,554
PRIOR FILING DATE: 1997-07-03
NUMBER OF SEQ ID NOS: 44
SEQ ID NO 16
LENGTH: 451
TYPE: PRT
ORGANISM: Artificial
FEATURE:
NAME/KEY: Artificial
LOCATION: 1-451
OTHER INFORMATION: Heavy chain sequence derived from MAE11
US-09-109-207C-16

Query Match      43.1%; Score 1223; DB 3; Length 451;
Best Local Similarity 64.1%; Pred. No. 1.2e-101;
Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;

QY 188 ENLDLTLQENSILYTIPIKGF---GSHLLPFAFLHGNPWLNCETILYFRRWLQDAENV 243
Db 74 DSKNTFYLOMNSLRADETAVYYCARGSHY---FGH---W-----HPAVWGQGLTVTV 119

QY 244 -----YVMKQVDVKAVTNSVASVQCDNSDKFP-----VYKYPGKGC 280
Db 120 SSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA--- 176

QY 281 PTIGDEG-----DTDLY-----DYYPEDETEGDKYR-----HTCPQCPA 315
Db 177 -VLQSSGLYSLSSVWTVSPSSSLGTQYICNVNHPKPSNTKYDKKVEPKSCDKTHTCPQCPA 235

QY 316 PEALGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 375
Db 236 PELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 295

QY 376 REEQYNSTYRVWSVLTVLHODWMLNGKEYKCKVSNKALPVPKEITISKAKGQPREPOVYTL 435
Db 296 REEQYNSTYRVWSVLTVLHODWMLNGKEYKCKVSNKALPVPKEITISKAKGQPREPOVYTL 355

QY 436 PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLT 495
Db 356 PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLT 415

QY 496 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 531
Db 416 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 451
```

Search completed: October 14, 2003, 06:32:12  
Job time : 24.734 secs